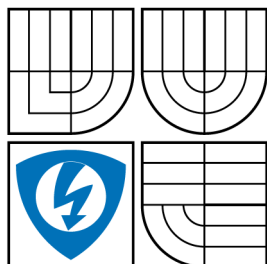


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ  
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY  
A KOMUNIKAČNÍCH TECHNOLOGIÍ  
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND  
COMMUNICATION  
DEPARTMENT OF TELECOMMUNICATIONS

# ANALÝZA FONACE U PACIENTŮ S PARKINSONOVOU NEMOCÍ

ANALYSIS OF PHONATION IN PATIENTS WITH PARKINSON'S DISEASE

DIPLOMOVÁ PRÁCE  
MASTER'S THESIS

AUTOR PRÁCE  
AUTHOR

Bc. TOMÁŠ KOPŘIVA

VEDOUCÍ PRÁCE  
SUPERVISOR

Ing. JIŘÍ MEKYSKA, Ph.D.

BRNO 2015



**VYSOKÉ UČENÍ  
TECHNICKÉ V BRNĚ**

**Fakulta elektrotechniky  
a komunikačních technologií**

**Ústav telekomunikací**

# Diplomová práce

magisterský navazující studijní obor  
**Telekomunikační a informační technika**

**Student:** Bc. Tomáš Kopřiva

**ID:** 136542

**Ročník:** 2

**Akademický rok:** 2014/2015

## NÁZEV TÉMATU:

### **Analýza fonace u pacientů s Parkinsonovou nemocí**

## POKYNY PRO VYPRACOVÁNÍ:

U 90 % pacientů s Parkinsonovou nemocí se vyskytuje porucha motorické realizace řeči, která se označuje hypokinetická dysartrie. Tento druh dysartrie se projevuje především v oblastech fonace, artikulace, prozodie, plynulosti řeči a faciokineze. Především z hlediska fonace je řeč monotónní, u prodloužených samohlásek dochází ke kolísání kmitočtu základního tónu, pacienti nejsou schopni udržet stejnou intenzitu, projevuje se hlasový tremor, atd. V rámci této práce bude analyzována skupina pěti českých samohlásek vyslovená jak pacienty s Parkinsonovou nemocí, tak zdravými kontrolami. Pomocí akustické analýzy budou vybrány ty samohlásky, ve kterých se hypokinetická dysartrie nejvíce projevuje, a následně bude otestován jejich význam při klasifikaci tohoto onemocnění.

## DOPORUČENÁ LITERATURA:

- [1] RUSZ, J.; CMEJLA, R.; RUZICKOVA, H.; RUZICKA, E.: Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. J Acoust Soc Am, roč. 129, č. 1, 2011: s. 350–367.
- [2] TSANAS, A.; LITTLE, M.; MCSHARRY, P.; RAMIG, L.: Accurate Telemonitoring of Parkinson's Disease Progression by Noninvasive Speech Tests. IEEE T Bio-Med Eng, roč. 57, č. 4, 2010: s. 884–893.
- [3] TSANAS, A.; LITTLE, M. A.; MCSHARRY, P. E.; RAMIG, L. O.: Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. J R Soc Interface, roč. 8, č. 59, 2010: s. 842–855.

**Termín zadání:** 9.2.2015

**Termín odevzdání:** 26.5.2015

**Vedoucí práce:** Ing. Jiří Mekyska, Ph.D.

**Konzultanti diplomové práce:**

**doc. Ing. Jiří Mišurec, CSc.**  
Předseda oborové rady

## ABSTRAKT

Tato práce se zabývá analýzou fonace u pacientů s Parkinsonovou nemocí (PN). Přibližně 90 % pacientů s Parkinsonovou nemocí trpí motorickou poruchou řeči nazývanou hypokinetická dysartrie. Je navržen systém pro rozpoznání Parkinsonovy nemoci z řečových signálů a je otestováno několik typů příznaků. Pro klasifikaci je použita česká řečová databáze pacientů s Parkinsonovou nemocí PARCZ. Obsahuje 84 pacientů s PN a 49 zdravých kontrolních řečníků. Výsledky jsou vyhodnoceny dvěma způsoby. Nejprve jsou příznaky analyzovány jednotlivě Spearmanovým korelačním testem, vzájemnou informací a Mann-Whitneyho U testem. Klasifikace je založena na náhodných stromech společně s validací leave-one-out. V druhém kroku je použit algoritmus výběru příznaků SFFS pro dosažení co nejlepšího výsledku klasifikace. Navržený systém je otestován pro jednotlivá pohlaví zvlášť i dohromady. Nejlepší výsledek pro obě pohlaví dohromady vyjadřuje přesnost klasifikace 89,47 %, sensitivita 91,67 % a specifická 85,71 %. Výsledky práce ukázaly, že pro analýzu fonace mají největší význam dlouhé realizace vokálů vyslovené s maximální nebo minimální intenzitou (ne šeptem).

## KLÍČOVÁ SLOVA

Hypokinetická dysartrie, Parkinsonova nemoc, řečové příznaky, výběr příznaků, SFFS, analýza fonace.

## ABSTRACT

This work deals with analysis of phonation in patients with Parkinson's disease (PD). Approximately 90 % of patients with Parkinson's disease suffer from speech motor dysfunction called hypokinetic dysarthria. System for Parkinson's disease analysis from speech signals is proposed and several types of features are examined. Czech Parkinson's speech database called PARCZ is used for classification. This dataset consists of 84 PD patients and 49 healthy controls. Results are evaluated in two ways. Firstly, features are individually analysed by Spearman correlation, mutual information and Mann-Whitney U test. Classification is based on random forests along with leave-one-out validation. Secondly, SFFS algorithm is employed for feature selection in order to get the best classification result. Proposed system is tested for each gender individually and both genders together as well. Best result for both genders together is expressed by accuracy 89,47 %, sensitivity 91,67 % and specificity 85,71 %. Results of this work showed that the most important vowel realizations for phonation analysis are sustained vowels pronounced with maximum or minimum intensity (not whispering).

## KEYWORDS

Hypokinetic dysarthria, Parkinson's disease, features, feature selection, SFFS, phonation analysis.

KOPŘIVA, Tomáš *Analýza fonace u pacientů s Parkinsonovou nemocí*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2015. 62 s. Vedoucí práce byl Ing. Jiří Mekyska, Ph.D.

## PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Analýza fonace u pacientů s Parkinsonovou nemocí“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....

(podpis autora)

## PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce Ing. Jiřímu Mekyskovi, Ph.D. za odborné vedení a cenné rady při psaní práce. Mé poděkování také patří rodičům za neustálou podporu během celého studia.

Brno .....

.....

(podpis autora)

## PODĚKOVÁNÍ

Výzkum popsáný v této diplomové práci byl realizován v laboratořích podpořených z projektu SIX; registrační číslo CZ.1.05/2.1.00/03.0072, operační program Výzkum a vývoj pro inovace.

Brno .....

.....  
(podpis autora)

# OBSAH

Úvod	9
<b>1 Parkinsonova nemoc</b>	
<b>a hypokinetická dysartrie</b>	<b>10</b>
1.1 Parkinsonova nemoc . . . . .	10
1.2 Hypokinetická dysartrie . . . . .	10
1.2.1 Fonace . . . . .	10
1.2.2 Artikulace . . . . .	11
1.2.3 Prozodie . . . . .	11
1.2.4 Plynulost řeči . . . . .	11
1.2.5 Faciokineze . . . . .	12
<b>2 Řečové příznaky</b>	<b>13</b>
2.1 Popis fonace . . . . .	13
2.2 Popis kvality hlasu . . . . .	14
2.3 Popis kmitání hlasivek . . . . .	14
2.4 Nelineární dynamické parametry . . . . .	16
<b>3 Rozvoj problematiky</b>	<b>20</b>
3.1 Přehled publikací . . . . .	20
3.2 Shrnutí . . . . .	22
<b>4 Popis systému</b>	<b>25</b>
4.1 Databáze PARCZ . . . . .	25
4.2 Výpočet řečových příznaků . . . . .	26
4.3 Statistická analýza . . . . .	30
4.3.1 Spearmanův test pořadové korelace . . . . .	30
4.3.2 Vzájemná informace . . . . .	31
4.3.3 Mann-Whitneyho U test . . . . .	32
4.4 Výběr příznaků . . . . .	32
4.4.1 Metody výběru . . . . .	32
4.4.2 Sekvenční dopředná selekce příznaků . . . . .	33
4.5 Klasifikace . . . . .	35
4.5.1 Klasifikační stromy . . . . .	35
4.5.2 Validace a proces klasifikace . . . . .	35

<b>5</b>	<b>Výsledky</b>	<b>37</b>
5.1	Individuální analýza příznaků . . . . .	37
5.2	Kombinace řečových příznaků . . . . .	43
<b>6</b>	<b>Závěr</b>	<b>49</b>
	<b>Literatura</b>	<b>51</b>
	<b>Seznam symbolů, veličin a zkratk</b>	<b>55</b>
	<b>Seznam příloh</b>	<b>60</b>
<b>A</b>	<b>Obsah CD</b>	<b>61</b>
<b>B</b>	<b>Obsah adresáře feature_toolbox</b>	<b>62</b>



# ÚVOD

Parkinsonova nemoc je druhé nejrozšířenější neurodegenerativní onemocnění, které postihuje 100 až 150 pacientů na 100 000 obyvatel. Byla popsána již roku 1817 lékařem Jamesem Parkinsonem. Diplomová práce se zabývá analýzou fonace pacientů s Parkinsonovou nemocí, u kterých se ve vysoké míře (90 %) vyskytuje motorická porucha řeči nazývaná hypokinetická dysartrie.

Cílem této práce je navrhnout systém pro klasifikaci, výběr vhodných samohlásek a řečových příznaků, ze kterých jsou projevy hypokinetické dysartrie nejlépe patrné. Význam zvolených samohlásek, příznaků a různých metod pro dosažení co nejvyšší přesnosti klasifikace je otestován na řečovém korpusu PARCZ, který je zaznamenáván ve Fakultní nemocnici u sv. Anny v Brně.

První kapitola se zabývá Parkinsonovou nemocí, hypokinetickou dysartrií a oblastem, ve kterých lze poruchy řeči sledovat. Další část popisuje řečové příznaky používané pro popis patologického hlasu a uvádí jejich možné rozdělení do několika skupin. Přehled vybraných publikací o současném stavu problematiky uvádí třetí kapitola. Stěžejní část práce představuje čtvrtá kapitola, ve které je kompletně popsán navržený systém a způsob jeho implementace v prostředí Matlab. V závěrečné kapitole jsou uvedeny dosažené výsledky a jejich analýza.

# 1 PARKINSONOVA NEMOC

## A HYPOKINETICKÁ DYSARTRIE

Kapitola se věnuje Parkinsonově nemoci (PN), která je nejčastější příčinou vzniku motorické poruchy řeči označované jako hypokinetické dysartrie (HD). Dále je popsán vliv HD na jednotlivé složky vytvářející řeč.

### 1.1 Parkinsonova nemoc

Parkinsonova nemoc se řadí do skupiny neurodegenerativních onemocnění zasahující centrální nervovou soustavu. Výsledkem je progresivní úbytek dopaminergních neuronů především v části mozku zvané substantia nigra (černá substance). Tyto neurony jsou zodpovědné za tvorbu dopaminu, který slouží jako neuropřenašeč umožňující přenos impulsů mezi nervovými buňkami. Mezi hybné symptomy PN [27] patří např. bradykineze, rigidita, třes a v oblasti motorické realizaci řeči u PN lze poruchy souhrnně označit pojmem hypokinetická dysartrie.

### 1.2 Hypokinetická dysartrie

Dysartrií nazýváme poruchu pohybové složky tvorby řeči [39] a patří mezi nejrozšířenější neurogenní poruchy komunikace [17].

Hypokinetická dysartrie je nejčastější porucha motorické stránky řeči u Parkinsonovy nemoci. V klinické praxi se člení dle závažnosti na lehký, střední a těžký stupeň [39]. V nejtěžších případech, kdy je ztracena schopnost artikulace, se stav nazývá anartrie. Velká část pacientů (70–90 %, [30, 21]) postižená Parkinsonovou nemocí má kvůli HD problémy s řečí a polykáním nebo se problémy vyskytnou v průběhu onemocnění. Deficity v řečové oblasti ovlivňují velkou měrou život pacienta – omezení schopnosti komunikace, zhoršení emočního a psychologického stavu má nepříznivý vliv na celkovou kvalitu života.

Hypokinetická dysartrie negativně působí na dílčí části vytvářející řeč – tzn. oblast fonace, artikulace, prozodie, plynulosti řeči a faciokineze [21]. Hlavním cílem práce je analýza fonace, stručně jsou popsány i ostatní zmíněné oblasti.

#### 1.2.1 Fonace

Fonace označuje proces vytváření hlasu. V této oblasti se poruchy projevují nejdříve. Důsledkem svalové rigidity se na činnosti hlasivek projevuje hlasový tremor (třes), nedokonalé sevření nebo pomalé chvění hlasivek. Řeč pacientů při vyslovování delších

promluvy je vnímána jako tichá a monotónní. Pacienti trpící PN mají také problém se zesílením nebo zeslabením hlasu a intenzita řeči se během promluvy v čase snižuje. Kvalita hlasu je negativně ovlivněna hypofonií a dysfonií. Hypofonie se projevuje sníženou intenzitou hlasu, jeho tichostí a nevýrazností. Dysfonie značí degradaci kvality hlasu – drsnost a chraptivost [39].

Důležitým parametrem je kmitočet základního tónu  $F_0$ , jehož střední hodnota je zvýšená nejen u delších promluv, ale také u prodloužených vokálů. Bylo zjištěno [22], že střední hodnota  $F_0$  narůstá s rozvojem nemoci a tím pádem lze tento parametr využít k určení progresu PN [21].

Respirace neboli dýchání úzce souvisí s oblastí fonace. Symptomy bradykineze a rigidita mají vliv na svalstvo při nádechu (inspirium) a výdechu (expirium). Lze sledovat zkrácení fází dýchání, neefektivní hospodaření s dechem nebo náhlé změny tempa řeči. Pacienti s PN obtížně koordinují dýchání a mluvení, dochází k opakování počátečních slabik, nevhodnému umístění pauz a nádechů nebo hypernazalitě [39, 21].

### 1.2.2 Artikulace

Symptomy Parkinsonovy nemoci (bradykineze, hypokineze atd.) se podobně jako u fonace projevují také na svalstvu orofaciální oblasti. Jsou zasaženy aktivní artikulační orgány [21]. U artikulace dochází ke špatné výslovnosti souhlásek, výslovnost samohlásek narušena není [39]. Diadochokinetické úlohy (DDK – Diadochokinetic Tasks) slouží pro posouzení přesnosti artikulace. Ty jsou založeny na principu rychlého opakování slabik tvořených kombinací souhláska-samohláska (např. pa-ta-ka-pa-ta-ka ...). Vlivem špatné artikulace řeč působí nedbalým dojmem, je špatně srozumitelná s nepřesnou výslovností. Tyto poruchy se objevují na úrovni slov, vět i delších celků [39].

### 1.2.3 Prozodie

Základní parametry prozodie jsou vyjádřeny: kmitočtem  $F_0$  (popisuje výšku hlasu), hlasitostí a tempem řeči. Mezi další prozodické rysy lze zařadit např. rytmus, melodii, přízvuk [33]. Význam prozodie v komunikaci je značný. Intonací na konci věty určujeme její typ nebo jinou hlasitostí či výškou hlasu zdůrazňujeme jednotlivá slova nebo části věty.

### 1.2.4 Plynulost řeči

U pacientů se vyskytuje opakování jednotlivých slabik nebo slov (palilalie) a vytváření neúmyslných pauz během řečového projevu (hezitace). Plynulost řeči souvisí

s prozodií – projevuje se zpomalení (bradyfemie) nebo zrychlení (tachyfemie) tempa řeči [21].

### **1.2.5 Faciokineze**

Tato oblast v sobě zahrnuje vliv orofaciálního svalstva na artikulaci, žvýkání, polykání a mimiku [39]. U pacientů se vyskytuje hypomimie, která vede k omezené schopnosti vyjádřit emoce, měnit výraz obličeje a znesnadňuje tak komunikaci v běžném životě.

## 2 ŘEČOVÉ PŘÍZNAKY

Parametrizace je postup, při němž jsou z řečového signálu extrahovány řečové příznaky, které popisují informaci obsaženou v signálu formou skaláru, vektoru nebo matice. Vhodná volba řečových příznaků je prvním krokem pro správnou funkčnost veškerých klasifikačních systémů. Kapitola se zabývá jejich popisem a členění je do několika skupin. Přehled popsanych příznaků je uveden v tab. 2.4.

### 2.1 Popis fonace

Frekvence základního tónu  $F_0$  patří mezi základní parametry pro popis řečového signálu. Jedná se o frekvenci, na které kmitají hlasivky a odpovídá výšce hlasu. K určení diagnózy PN se často používá její střední hodnota  $\text{mean}(F_0)$  a odchylka  $\text{std}(F_0)$ . Dále lze také sledovat rozdíl maximálních a minimálních hodnot v řeči označený jako  $F_0$  Variation range:  $F_0\text{VR} = F_{0\text{max}} - F_{0\text{min}}$  (Hz). Mezi další parametry patří relativní směrodatná odchylka a relativní kolísání. Jsou vypočítány normalizací směrodatné odchylky a  $F_0\text{VR}$  [21]:

$$\text{rel}F_0\text{SD} = \frac{\text{std}(F_0)}{\text{mean}(F_0)} \cdot 100 \text{ (}\%) \quad \text{a} \quad \text{rel}F_0\text{VR} = \frac{F_0\text{VR}}{\text{mean}(F_0)} \cdot 100 \text{ (}\%). \quad (2.1)$$

Jitter popisuje kolísání  $F_0$  v čase tzn. vyjadřuje třes a chvění hlasivek. Vypočítá se jako průměrný rozdíl délek dvou sousedních period hlasivkových pulsů vydělený průměrnou délkou periody. Je vyjádřen v procentech. Obvyklé hodnoty se při normální řeči pohybují kolem 0,5–1 %, změna je patrná při hodnotě aspoň 2 % [25]. Mezi jitter se dále řadí příznaky RAP (Relative Average Perturbation) a PPQ5 (five-point Pitch Perturbation Quotient). RAP se vypočítá jako rozdíl periody a průměru vybrané periody a jejích sousedních period dělený průměrnou délkou periody. PPQ5 je výpočtově shodný jako RAP, odlišný je pouze průměr, který se počítá z dané periody a dvou sousedních period z každé strany (tj. celkově čtyři sousední periody).

Pro odstranění závislosti na pohlaví byl navržen řečový příznak PPE (Pitch Period Entropy) [18]. PPE měří narušenou schopnost udržet stabilitu  $F_0$  během procesu tvorby hlasu.

Shimmer je definován jako rozdíl velikostí dvou amplitud sousedních hlasivkových pulsů dělený jejich průměrnou velikostí. Vyjadřuje kolísání neboli vibraci velikosti pulsů a udává se v decibelech. Pro normální řeč je shimmer v rozmezí 0,4–0,21 dB, kolísání je slyšitelné od 1 dB [25]. Mezi druhy shimmeru patří např. APQ3 (three-point Amplitude Perturbation Quotient), APQ5 (five-point Amplitude Perturbation Quotient) a APQ11 (11-point Amplitude Perturbation Quotient) [9]. Při výpočtu se používají velikosti pulsů, jinak je postup výpočtu shodný s jitterem. Mezi hlasivkové

kvocienty GQ (Glottal Quotients) patří směrodatná odchylka doby kdy jsou hlasivky otevřené ( $GQ_{\text{open}}$ ) a uzavřené ( $GQ_{\text{closed}}$ ).

Intenzita neboli hlasitost je jedna ze základních prozodických vlastností. Lze ji vyjádřit krátkodobou energií signálu  $E$  nebo TKEO (Teager-Kaiser Energy Operator). Kolísání intenzity lze popsat směrodatnou odchylkou, shimmerem, rozdílem maximální a minimální hodnoty energie  $\Delta E$  a rozdílem maximální a minimální amplitudy hlasivkových pulsů  $\Delta A$ .

## 2.2 Popis kvality hlasu

Jednoduchým příznakem pro hodnocení kvality hlasu je počet průchodů signálu nulovou úrovní ZCR (Zero-Crossing Rate). Obdobný příznak HZ-CRR (High Zero-Crossing Rate Ratio) je definován jako počet segmentů, jejichž ZCR je větší než hodnota 1,5-násobku průměrného počtu průchodů nulovou úrovní signálu [35]. Pro hodnocení kvality lze také použít podíl neznělých segmentů FLUF (Fraction of Locally Unvoiced Frames) [1].

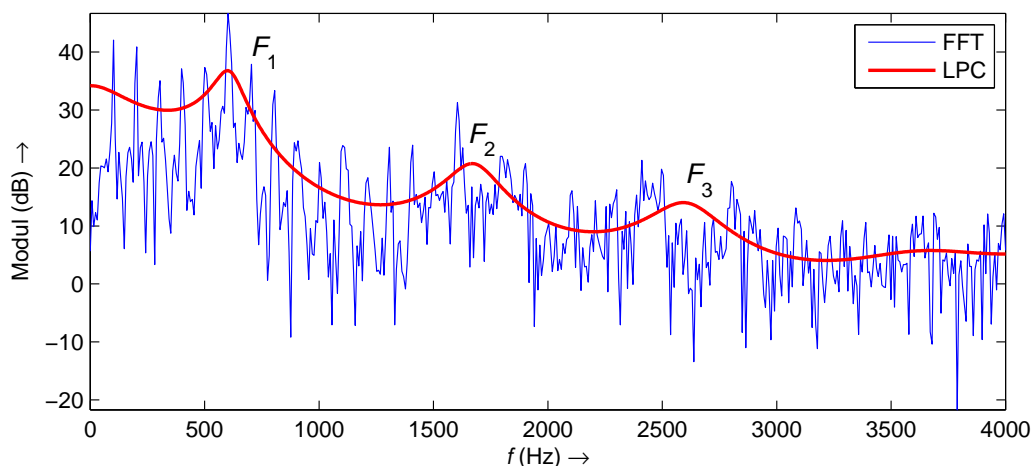
Spektrální tok SF (Spectral Flux) se vypočítá jako rozdíl výkonového spektra mezi dvěma sousedními segmenty [29]. Dále SDBM (Spectral Distance Based on Module) je suma rozdílů modulů sousedních segmentů [1] a SDBP (Spectral Distance Based on Phase) je suma rozdílů fází sousedních segmentů [1].

Mezi příznaky založené na reálném kepru patří CPP (Cepstral Peak Prominence). Jedná se o rozdíl velikostí (vyjádřený v dB) mezi maximem první rahmonické a odpovídající hodnotě na přímce regrese, která se nachází přímo pod daným vrcholem [13]. Dále se také používají PECM (Pitch Energy Cepstral Measure) [1] a RR (Rahmonic Ratio) [14].

Příznaky pro odhad úrovně šumu v signálu jsou např. HNR (Harmonic-to-Noise Ratio) [20], NHR (Noise-to-Harmonic Ratio) [20], SPI (Soft Phonation Index) [8] nebo VTI (Voice Turbulence Index) [8].

## 2.3 Popis kmitání hlasivek

Při tvorbě řeči prochází proud vzduchu z plic přes hlasivky do nadhrtanových dutin (hrdelní, nosní a ústní). Kvaziperiodický budící signál vycházející z hlasivek se označuje jako základní hlasový tón. Při jeho průchodu nadhrtanovými dutinami v nich dochází k rezonancím na kmitočtech, které se nazývají formanty ( $F_1, F_2, \dots, F_n$ ). Znázornění formantů pro vokál [e] zobrazuje obr. 2.1. Jsou nejvýznamnější akustickou charakteristikou samohlásek [25]. Na hodnoty formantů má vliv postavení



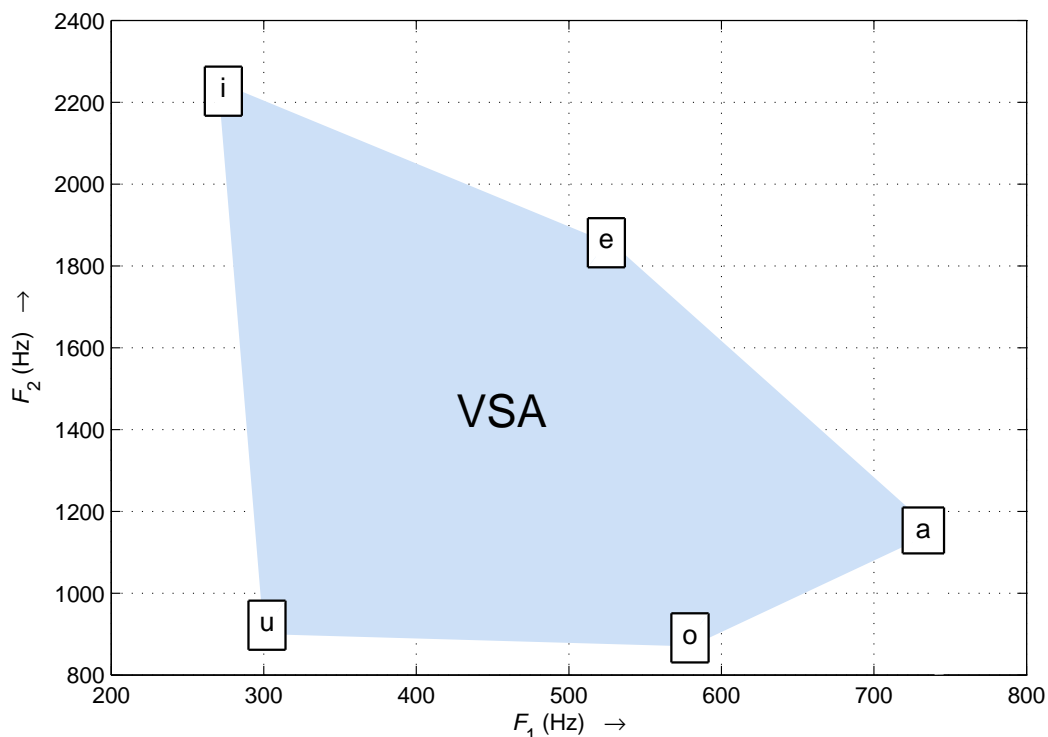
Obr. 2.1: Formanty samohlásky [e]

jazyka, rtů, zubů a měkkého patra, kterými se mění objem a uspořádání příslušných dutin. Vzhledem k jejich vlastnostem se využívají např. ke sledování hybnosti jazyka a popisu vokálů pacientů s PN. Jejich nevýhodou je silná závislost na věku a pohlaví. Proto je používán poměr  $F_{2i}/F_{2u}$  (odpovídá samohláskám [i] a [u]), který dokáže odlišit zdravou řeč od řeči s HD bez závislosti na pohlaví [21]. U samohlásek je nejvýznamnější artikulační charakteristika poloha a pohyb jazyka. V horizontálním směru pohybu jazyka se rozlišují samohlásky přední, střední a zadní. Ve vertikálním směru jsou rozděleny na vysoké, středové a nízké samohlásky. Tuto klasifikaci poprvé popsal Ch. F. Hellwag a závislost se zobrazuje ve formě vokalického trojúhelníku (tzv. Hellwagův trojúhelník [25]) – viz tab. 2.1. Byla také dokázána souvislost tohoto rozdělení s formantovými kmitočty. Horizontální pohyb ovlivňuje velikost druhého formantu a ve vertikálním směru se mění první formant [10].

Z vokalického trojúhelníku vychází příznak VSA (Vowel Space Area). Prostřednictvím formantových kmitočtů vyjadřuje míru pohybu jazyka. Z řeči jsou extrahovány vokály [a], [e], [i], [o], [u] a poté se odhadnou hodnoty prvních dvou formantů. Hodnoty  $F_1$  a  $F_2$  v grafu vytvoří pětiúhelník, jehož obsah je roven hodnotě

Tab. 2.1: Vokalický trojúhelník

		vodorovný pohyb jazyka		
		přední	střední	zadní
$\leftarrow F_1$	vysoké	[i]		[u]
	středové		[e]	[o]
	nízké		[a]	
		$\leftarrow F_2$		



Obr. 2.2: Vokální oblast hlasového traktu

VSA (obr. 2.2). U příznaku lnVSA jsou hodnoty formantů nejdříve logaritmovány přirozeným logaritmem a následně je proveden výpočet stejným způsobem jako u VSA [31]. Pro potlačení závislosti VSA na pohlaví byl navržen příznak FCR (Formant Centralization Ratio) [31]:

$$\text{FCR} = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}}, \quad (2.2)$$

kde  $F_{2u}$  je frekvence druhého formantu samohlásky [u] atd. Tím, že je FCR vyjádřen jako zlomek, se provádí určitá normalizace a zvyšuje se tak robustnost příznaku.

Další příznak nezávislý na pohlaví VAI (Vowel Articulation Index) [32] je definován jako převrácená hodnota FCR:  $\text{VAI} = 1/\text{FCR}$ .

## 2.4 Nelineární dynamické parametry

V patologické řeči se vyskytují nelineární jevy (aperiodicity v čase, sub-harmonické složky, degradace signálu), které nelze popsat běžnými metodami zpracování signálů – tj. příznaky vypočítanými z časového průběhu, spektra apod. Ty převážně vycházejí z lineárního modelu tvorby řeči a z tohoto důvodu jimi nelze vhodně



popsat kvalitativně velmi degradovaný řečový signál. Je proto nutné pro získání užitečné informace ze signálu použít příznaky založené na nelineární dynamické analýze [38, 11].

Korelační dimenze CD (Correlation Dimension) statisticky měří geometrii shluku bodů generovaných ve fázovém prostoru. Vztahuje se k počtu proměnných nutných pro generování odpovídajícího procesu. Čím větší komplexity systém dosahuje, tím vyšší má dimenzi a tím pádem je nutné více proměnných pro popis dynamiky systému [38]. Z tohoto důvodu může být vhodným parametrem pro popis komplexnosti systému. Fraktální dimenze FD (Fractal Dimension) představuje část dimenze geometrického objektu. FD je relativní míra počtu základních stavebních bloků, které vytváří vzor. Odhad fraktální dimenze je možný buď přímo v časové oblasti nebo ve fázovém prostoru [38]. Příznak ZL (Ziv-Lempel) měří komplexnost systému tím, že kvantifikuje pravidelnosti uvnitř časové posloupnosti. Signál se nejprve transformuje na binární posloupnost a následně je analyzován – jsou nalezeny „vzory“, které se v signálu opakují [2, 38]. Dlouhodobé závislosti v signálu mohou být odhadnuty Hurstovým exponentem HE (Hurst Exponent) [2].

Entropie obecně udává stupeň neurčitosti náhodné proměnné a je maximální pro nekorelovaný náhodný signál. Nárůst entropie tak obvykle odpovídá nárůstu komplexnosti systému [38]. Patří zde např. Shannonova entropie SE [11], Rényiho entropie RE [11], aproximační entropie AE [38], RPDE (Recurrence Probability Density Entropy) [20] atd.

Do kategorie nelineárních dynamických parametrů se také řadí minimum funkce vzájemné informace FMMI (First Minimum of Mutual Information function). Ta měří vzájemnou závislost mezi dvěma proměnnými [11]. Dále největší Lyapunovův exponent LLE (Largest Lyapunov Exponent) odpovídá citlivosti systému na počáteční podmínky [38].

Tab. 2.2: Přehled řečových příznaků

Popis fonace	
$F_0$	kmitočet základního tónu
$\text{mean}(F_0)$	střední hodnota $F_0$
$\text{std}(F_0)$	směrodatná odchylka $F_0$
jitter	kolísání $F_0$
$F_0\text{VR}$	rozdíl maximálních a minimálních hodnot $F_0$
$\text{rel}F_0\text{SD}$	relativní směrodatná odchylka $F_0$
$\text{rel}F_0\text{VR}$	relativní kolísání
RAP	Relative Average Perturbation
PPQ	five-point Pitch Perturbation Quotient
PPE	Pitch Period Entropy
shimmer	kolísání $E$ nebo $A$
APQ3	three-point Amplitude Perturbation Quotient
APQ5	five-point Amplitude Perturbation Quotient
APQ11	11-point Amplitude Perturbation Quotient
$\text{GQ}_{\text{open}}$	směrodatná odchylka doby, kdy jsou hlasivky otevřené
$\text{GQ}_{\text{closed}}$	směrodatná odchylka doby, kdy jsou hlasivky uzavřené
$E$	krátkodobá energie signálu
TKEO	Teagrův-Kaiserův energetický operátor
$\Delta E$	rozdíl maximální a minimální hodnoty $E$
$\Delta A$	rozdíl maximální a minimální hodnoty $A$
Popis kvality hlasu	
ZCR	počet průchodů signálu nulovou úrovní
HZ-CRR	High Zero-Crossing Rate Ratio
FLUF	podíl neznělých segmentů
SF	spektrální tok
SDBM	změna spektra modulů
SDBP	změna spektra fází
CPP	Cepstral Peak Prominence
PECM	Pitch Energy Cepstral Measure
RR	Rahmonic Ratio
Odhad úrovně šumu	
HNR	Harmonic-to-Noise Ratio
NHR	Noise-to-Harmonic Ratio
SPI	Soft Phonation Index
VTI	Voice Turbulence Index

Popis kmitání hlasivek	
$F_n$	formanty
$F_{2i}/F_{2u}$	poměr formantů – nezávislý na pohlaví
VSA	oblast hlasového traktu
lnVSA	logaritmická oblast hlasového traktu
FCR	centralizační poměr formantů
VAI	artikulační index samohlásek
Nelineární dynamické parametry	
CD	korelační dimenze
FD	fraktální dimenze
ZL	komplexnost systému Ziv-Lempel
HE	Hurstův exponent (dlouhodobé závislosti signálu)
SE	Shannonova entropie
RE	Rényiho entropie
AE	aproximační entropie
RPDE	Recurrence Probability Density Entropy
FMFI	minimum funkce vzájemné informace
LLE	největší Lyapunovův exponent

## 3 ROZVOJ PROBLEMATIKY

Kapitola uvádí přehled vybraných publikací, které se zabývají fonací pro účely rozpoznání PN. Důraz je kladen na analýzu prodloužených vokálů. Stručně jsou popsány použité metody, databáze, příznaky, klasifikátory a dosažené výsledky. Kompletní přehled je uveden v tab. 3.1.

### 3.1 Přehled publikací

Práce autorů Belalcazar et al. [3] používá pouze čtyři příznaky založené na měření přítomného šumu: HNR, NNE (Normalized Noise Energy), CHNR (Cepstral HNR) a GNE (Glottal to Noise Excitation Ratio). Příznaky se počítají pro rámce a dané průběhy jsou následně popsány statistickými veličinami. Jedná se o střední hodnotu, směrodatnou odchylku, šikmost a špičatost. Použitá databáze se skládá z 50 mluvčích s Parkinsonovou nemocí a 50 zdravých kontrolních řečníků. Databáze je vyvážená věkově i pohlavím. U každého mluvčího jsou nahrány promluvy pěti španělských vokálů. Pro klasifikaci byl použit algoritmus  $k$ -nejbližších sousedů KNN. Pouze jeden typ příznaku poskytoval převážně špatné výsledky. Při použití všech příznaků a jejich transformace pomocí PCA (Principal Components Analysis) bylo dosaženo nejlepšího výsledku pro samohlásky [i] a [e] s přesností klasifikace 66,5 %. Naopak nejhorší výsledek (61,9 %) poskytuje vokál [u].

Španělský jazyk a stejný výběr samohlásek využívají i autoři [2] Orozco-Arroyave et al. Všechny 10 použitých příznaků je založeno na nelineární dynamické analýze. Jedná se o: CD, LLE, ZL, HE, RPDE, DFA (Detrended Fluctuation Analysis), AE, AE s Gaussovským jádrem (AEGK), Sample Entropy (SA) a SA s Gaussovským jádrem (SAGK). Statistické veličiny jsou zvoleny a vypočítány stejným způsobem jako ve výše zmíněné práci (tj. střední hodnota, směrodatná odchylka, šikmost a špičatost). Databáze se skládá z 20 pacientů s PN a 20 kontrolních řečníků. Pro klasifikaci se využívá SVM (Support Vector Machines). Nejprve byly analyzovány vokály samostatně. Nejlepšího výsledku 76,81 % opět dosáhla samohláska [i]. Při použití kombinace příznaků ze všech samohlásek nedošlo ke zlepšení (74,03 %). V tomto případě větší počet příznaků vede pouze ke zvýšení dimenzionality systému a nárůstu redundantní informace, která přesnost klasifikace nezvyšuje.

Hazan et al. [10] využívali dvě databáze – jednu v německém jazyce s 98 řečníky a další v anglickém jazyce s celkem 30 řečníky. Jejich rozdělení dle pohlaví a zdravotního stavu je uvedeno v tab. 3.1. Jako příznaky jsou použity první dva formanty vokálů [a], [u], [i]. Data pro klasifikaci byla reprezentována třemi různými způsoby: standardním způsobem, normalizovaně a logaritmicky. Klasifikace byla ve

všech případech provedena algoritmem podpůrných vektorů SVM s radiální bázo-  
vou funkcí jádra. Prvním scénářem bylo samostatné testování jednotlivých databází  
validací leave-one-out. V tomto případě bylo dosaženo přesnosti 93,94 % (optimální  
příznaky  $F_{2i}$ ,  $F_{2a}$ ,  $F_{1i}$ ) u americké databáze a 84,69 % (příznaky  $F_{2i}$ ,  $F_{2a}$ ,  $F_{2u}$ ) u ně-  
mecké. V dalším scénáři byl systém natrénován na americké databázi a otestovány  
byly záznamy německé databáze a následně byl postup obrácen. Přesnost klasifikace  
se v obou případech pohybovala kolem 75 %, jako nejvhodnější se ukázala logarit-  
mická reprezentace dat. V posledním scénáři byly databáze použity dohromady jak  
pro trénování, tak pro testování. Výsledkem jsou úspěšnosti 84,38 % (americká data-  
báze) a 79,25 % (německá databáze). Závěry publikace ukazují závislost optimálně  
vybraných příznaků na jazyku databází.

Autoři Rusz et al. [28] analyzovali vokály pacientů s PN v raném stadiu nemoci  
před nasazením dopaminerní léčby. Extrakce vokálů [a], [i], [u] byla provedena ze  
čtyř řečových cvičení: prodloužené vokály, opakování vět, čtení a monolog. Příznaky  
jsou založeny na formantech  $F_1$  a  $F_2$  určených ze zmíněných samohlásek. Autoři po-  
užili  $F_{2u}$ , poměr  $F_{2i}/F_{2u}$ , VSA a VAI. Určitým omezením práce je použitá databáze,  
která se skládá pouze z mužů – 20 pacientů s PN a 15 kontrolních řečníků. Pro  
hodnocení bylo použito nejprve několik statistických analýz (Pearsonův korelační  
koeficient, Kolmogorovův-Smirnovův test, ANOVA – Analysis of variance) a dále  
metoda klasifikace založená na teorému mini-max. Výsledky práce ukázaly nejnižší  
přesnost klasifikace 58,7 % u prodloužených vokálů. Naopak nejlepší průměrný vý-  
sledek 76,5 % dosáhlo řečové cvičení monologu spontánní řeči. V rámci testování  
monologu se ukázaly jako nejvhodnější příznaky VSA a  $F_{2i}/F_{2u}$  – přesnost 80,4 %  
resp. 80,0 %. V ostatních úlohách byly průměrné úspěšnosti klasifikace 73,5 % pro  
opakování vět a 65,5 % pro čtený text. Ve spontánní řeči je tedy ovlivnění artikulace  
vokálů dle autorů nejvíce zřejmé. Naopak prodloužené vokály jsou nejméně vhodné  
pro detekci artikulačních změn u pacientů v raném stadiu PN. V řeči se nedostatky  
projevily zvýšením formantu  $F_{2u}$ , snížením VSA, snížením rozdílu mezi vokály za-  
chycený  $F_{2i}/F_{2u}$  a abnormalitami v centralizaci formantů.

Další práce [18] autorů Little et al. se soustředí pouze na analýzu prodloužených  
vokálů. Používá nejen běžné metody výpočtu příznaků, ale také i příznaky založené  
na nelineární dynamické analýze. Z běžných metod je použito několik typů jitterů,  
shimmerů, NHR a HNR. U výpočtu nelineárních příznaků autoři využívají entropii  
RPDE, korelační dimenzi a DFA. Dále autoři navrhli nový příznak PPE, viz kap. 2.1.  
Po výpočtu řečových příznaků následuje jejich výběr neboli selekce. Tento krok je  
nutný z důvodu velkého množství možných kombinací příznaků a také vzhledem  
k jejich korelaci – např. hodnoty různých typů jitterů mohou být vzájemně korelo-  
vané, protože popisují obdobné vlastnosti signálu. Redukcí je vybrána kombinace 10  
nekorelovaných příznaků a jejich vektor slouží jako vstup klasifikátoru SVM s Gaus-

sovským radiálním jádrem. Následně jsou otestovány všechny možné kombinace 10 optimálně vybraných příznaků na databázi. Ta obsahuje 23 pacientů s PN, 8 kontrolních řečníků a celkem obsahuje 195 prodloužených vokálů. Nejvyšší přesnosti klasifikace 91,4 % dosáhla kombinace HNR, RPDE, DFA a PPE. Při použití všech 10 příznaků byla úspěšnost klasifikace 90,6 % a při nasazení jediného příznaku byl nejvhodnější PPE s 85,6 %. Výsledky práce ukázaly na vhodnost použití nelineárních dynamických metod, které vykazaly větší klasifikační přesnost (především jejich kombinace) než metody standardní. Také úspěšnost klasifikace neroste se zvyšujícím se počtem použitých příznaků, naopak optimální počet byl čtyři pro použitou množinu dat.

Nejvyšší přesnosti klasifikace dosáhli autoři [37] Tsanas et al. Pro výpočet příznaků použili prodloužený vokál [a]. Databáze se nazývá National Center for Voice and Speech a obsahuje celkem 263 prodloužených vokálů. Má 43 řečníků – 10 kontrolních a 33 s PN. Použité příznaky jsou vybrány z několika skupin: 1) jitter a shimmer včetně jejich variant, 2) HNR, DFA, GNE, VFER – Vocal Fold Excitation Ratio a EMD-ER – Empirical Mode Decomposition Excitation Ratio, 3) RPDE, PPE, GQ a příznaky vypočítané z kmitočtu základního tónu, 4) Melovské keprální koeficienty MFCC. Každý vokál z databáze je popsán celkem 132 příznaky a jejich koeficienty. Tento vysoký počet se následně redukuje. Autoři v práci porovnávají čtyři odlišné algoritmy pro jejich výběr – LASSO (Least Absolute Shrinkage and Selection Operator), mRMR (Minimum Redundancy Maximum Relevance), Relief a LLBFS (Local Learning-Based Feature Selection). Na rozdíl od jiných prací jsou testovány dvě klasifikační metody. Vyšší přesnosti klasifikace dosahoval algoritmus SVM s Gaussovským radiálním jádrem oproti náhodným stromům RF (Random Forests). Při využití kompletního počtu 132 příznaků je úspěšnost vysokých 97,7 %. Z důvodu možného ovlivnění výsledku způsobeného přetrénováním klasifikátoru autoři následně pracovali s výběrem 10 příznaků pro každý algoritmus. Redukcí počtu příznaků se přesnost zvýšila pro oba klasifikátory. Nejvyšší přesnost 98,6 % byla dosažena výběrem 10 příznaků metodou Relief v kombinaci s algoritmem SVM.

## 3.2 Shrnutí

Přehled použitých metod výše popsaných prací shrnuje tab.3.1. Je z ní patrné, že výsledky jednotlivých autorů jsou těžko porovnatelné mezi sebou. Každá práce totiž používá odlišný řečový korpus. A ty se liší v mnoha ohledech. Mají různý počet řečníků, odlišný poměr kontrolních řečníků a pacientů s PN, některé mají nevyvážený poměr mužů a žen. Dále se také často liší jazykem, nahranými vokály nebo tím, že pacienti v nich mají různý stupeň progresu PN.

Použité příznaky jsou vypočítány z formantů [10, 28], odhadnuty z šumu [3], z nelineární dynamické analýzy [2] nebo autoři využívají jejich různé kombinace [18, 37]. Pro výběr optimálních příznaků jsou využity různé metody, jenom Tsanas et al. se zabývali srovnáním více algoritmů a nejvyšší přesnosti klasifikace ve své práci dosáhli metodou Relief. Z uvedených prací pouze autoři Rusz et al. a Tsanas et al. používají vybrané statistické testy – konkrétně Kolmogorovův-Smirnovův test, ANOVA, Pearsonův korelační koeficient a vzájemnou informaci MI. U klasifikátorů převládá použití algoritmu SVM, který dosahuje vysoké procentuální úspěšnosti i nad 90 %.

Výsledky práce autorů Hazan et al. lze považovat za dostatečně robustní z důvodu nasazení velkého řečového korpusu v německém jazyce (celkem 98 řečníků) se stejným zastoupením mužů a žen. Analýzou vokálů a jejich formantů v kombinaci s klasifikátorem SVM dosáhli přesnosti 84,69 % [10]. Lepší výsledky publikovali další autoři (např. Little et al. – 91,4 % [18], Tsanas et al. – 98,6 % [37]), nicméně jimi použité řečové korpusy jsou mnohem menší (31 resp. 43 řečníků).

Tab. 3.1: Přehled publikací

Autoři	Kontrolní řečníci (Muži/Ženy)	Celkem kontrolní řečníci	Pacienti s PN (Muži/Ženy)	Celkem pacienti s PN	Databáze celkem	Jazyk	Samohlásky	Řečové příznaky	Klasifikátor	Přesnost
Belalcazar-Bolanos et al. (2013) [3]	-	50	-	50	100	španělština	[a], [e], [i], [o], [u]	HNR, NNE, CHNR, GNE	KNN	66,5 %
Hazan et al. (2012) [10]	7/7	14	8/8	16	30	angličtina	[a], [i], [u]	první a druhé formanty	SVM	93,94 %
Hazan et al. (2012) [10]	15/15	30	34/34	68	98	němčina	[a], [i], [u]	první a druhé formanty	SVM	84,69 %
Orozco-Arroyave et al. (2013) [2]	10/10	20	10/10	20	40	španělština	[a], [e], [i], [o], [u]	CD, LLE, ZL, HE, RPDE, DFA, AE, AEGK, SA, SAGK	SVM	76,81 %
Rusz et al. (2013) [28]	15/0	15	20/0	20	35	čeština	[i], [a], [u]	$F_{2u}$ , VSA, $F_{2i}/F_{2u}$ , VAI	metoda mini-max	80,4 %
Little et al. (2009) [18]	3/5	8	16/7	23	31	angličtina	[a]	HNR, RPDE, DFA, PPE, 5 typů jitteru, 6 typů shimmeru, NHR, CD	SVM	91,4 %
Tsanas et al. (2012) [37]	4/6	10	22/11	33	43	angličtina	[a]	jitter, shimmer a jejich varianty, RPDE, PPE, GQ, $F_0$ , HNR, NHR, DFA, GNE, VFER, EMD-ER, MFCC	SVM, RF	98,6 %



## 4 POPIS SYSTÉMU

Kapitola se zabývá navrhnutým systémem a jeho částmi. Nejprve je popsán řečový korpus PARCZ, na kterém byla klasifikace realizována. Dále jsou uvedeny jednotlivé bloky systému, jejich význam a způsob implementace v prostředí Matlab.

### 4.1 Databáze PARCZ

Prvním blokem systému na obr. 4.2 je řečový korpus PARCZ (Czech Parkinsonian Speech Database). Tato databáze řečových nahrávek byla poskytnuta Fakultní nemocnicí u sv. Anny v Brně. Její část s kompletně nahránými vokály, která je v práci použita, obsahuje 133 řečníků. Kontrolních řečníků je celkem 49 v poměru 24 žen a 25 mužů. Pacientů s PN obsahuje 84 – 36 žen a 48 mužů. Statistické charakteristiky databáze pro kontrolní řečníky a pacienty s PN jsou uvedeny v tab. 4.1. Každého řečníka popisuje kód NPxxx: N značí onemocnění – K je kontrolní řečník a P je pacient s Parkinsonovou nemocí; P je pohlaví pacienta – 1 pro ženu a 2 pro muže; xxx – pořadové číslo řečníka. Nahrávky jsou v jednobanálovém zvukovém formátu WAV s kódováním 16 bitů a vzorkovacím kmitočtem 16 kHz (jsou podvzorkovány z původních 48 kHz).

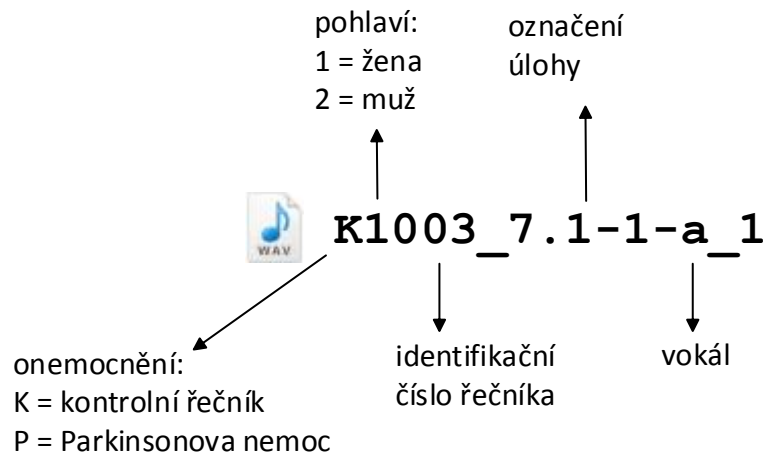
Tab. 4.1: Statistické parametry databáze PARCZ

	Kontrola		PN	
	Ženy	Muži	Ženy	Muži
Počet	24	25	36	48
Průměrný věk	62,15	65,58	68,80	66,31
std věku	9,31	8,73	7,66	8,62
Věkový rozsah	45–87	49–83	49–86	46–87

Databáze byla vytvořena na základě protokolu pro objektivní popis hypokinetické dysartrie, který obsahuje celkem 91 řečových úloh. Patří mezi ně např. monolog pacienta, čtení dlouhého textu, předříkávání vět náročných na artikulaci, recitace básně, diadochokinetické úlohy atd. Každá řečová úloha je označena číselným kódem. U části protokolu zaměřené na fonaci pacienti vyslovují vokály [a], [e], [i], [o], [u] čtyřmi způsoby. Nejprve krátce s přirozenou intenzitou (kód úlohy 7.1–1), dlouze s přirozenou intenzitou (7.1–2), dlouze s maximální možnou intenzitou (7.1–3) a dlouze s minimální intenzitou, ale ne šepem (7.1–4). Průměrné délky trvání a směrodatné odchylky (STD) v sekundách pro každou řečovou úlohu jsou uvedeny v tab. 4.2. Příklad úplného označení jedné nahrávky řečníka znázorňuje obr. 4.1.

Tab. 4.2: Průměrné délky a STD řečových úloh [s]

kód úlohy		Kontrola		PN	
		Ženy	Muži	Ženy	Muži
7.1-1	průměr	0,40	0,40	0,39	0,39
	std	0,09	0,11	0,11	0,10
7.1-2	průměr	0,76	0,73	0,78	0,73
	std	0,15	0,14	0,22	0,18
7.1-3	průměr	0,77	0,74	0,78	0,74
	std	0,16	0,13	0,33	0,20
7.1-4	průměr	0,59	0,57	0,59	0,61
	std	0,13	0,13	0,24	0,21

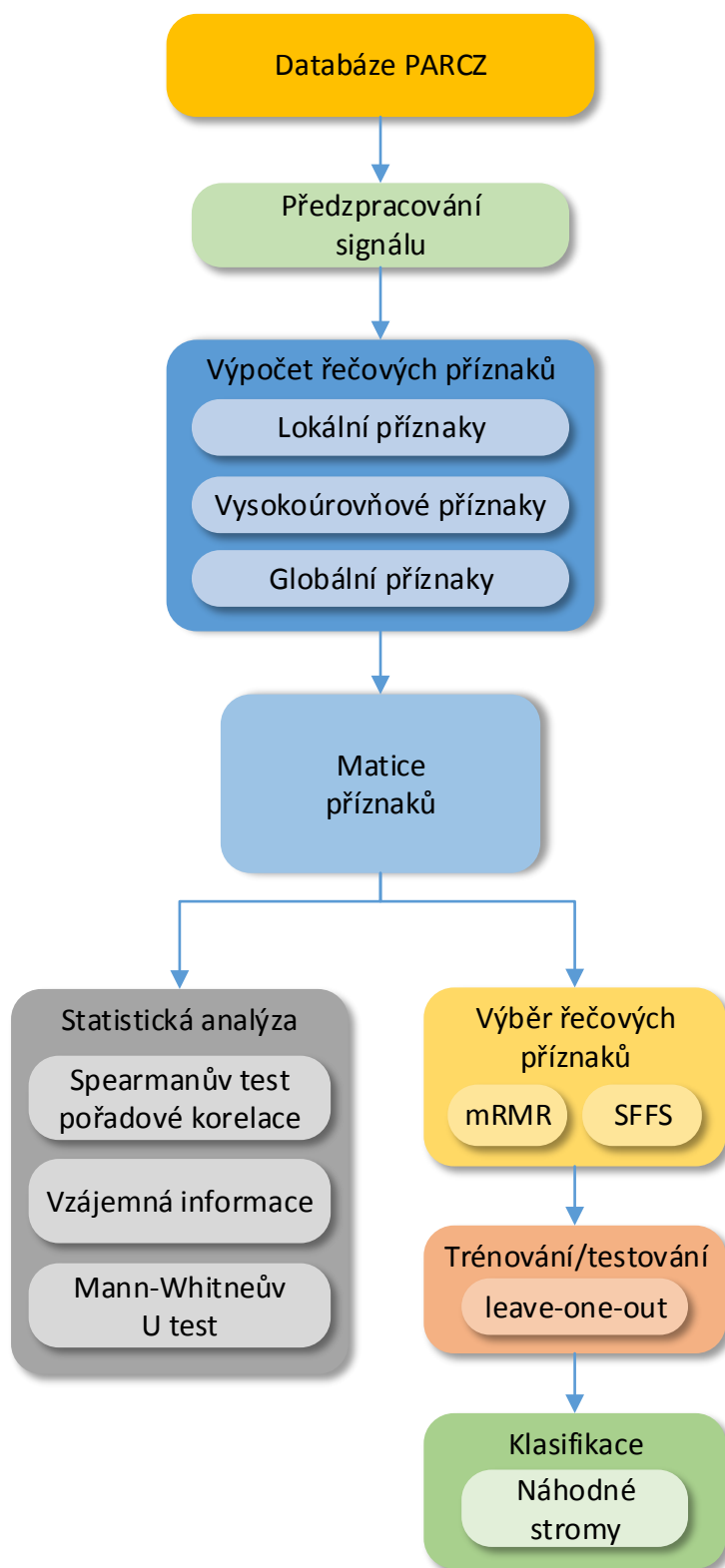


Obr. 4.1: Příklad pojmenování nahrávky

## 4.2 Výpočet řečových příznaků

Před extrakcí řečových příznaků je nejprve signál předzpracován – dojde k odstranění jeho stejnosměrné složky a normalizaci. U výpočtu některých příznaků (např.  $E$ , ZCR, fraktální dimenze atd.) je signál dále segmentován Hammingovým oknem s délkou 25 ms a překrytím 10 ms. Další použitou metodou je preemfáze, která se provádí filtrací řečového signálu horní propustí a tím dojde ke zvýraznění amplitud spektrálních složek vyšších kmitočtů. Preemfáze se provádí před výpočtem formantových kmitočtů.

Po předzpracování signálu následuje výpočet řečových příznaků, který je rozdělen do tří fází (obr. 4.2). Lokální příznaky počítané přímo ze signálu mají tvar



Obr. 4.2: Schéma navrhnutého systému

skaláru, vektoru nebo matic. Lokální příznaky ve formě vektorů a matic popisují změny v čase, protože se počítají pro každý segment samostatně. V tab. 4.4 je uveden přehled všech použitých lokálních příznaků včetně jejich typu – zda se jedná o skalární hodnotu nebo vektor.

V další fázi je provedena transformace lokálních příznaků ve formě vektorů (resp. matic) na skalární hodnoty – ty se nazývají vysokoúrovňové příznaky. Jsou implementovány základní statistické charakteristiky (tab. 4.5). Jedná se o střední hodnotu, směrodatnou odchylku, 1. a 99. percentil a mezipercentilový rozsah definovaný jako

$$ir = 99. \text{ percentil} - 1. \text{ percentil.} \quad (4.1)$$

Globální příznaky popsané v tab. 4.6 používají pro výpočet kombinaci řečových příznaků z různých signálů. Např. u VSA je nejprve nutné určit první dva formanty každého vokálu a pak teprve následuje vlastní výpočet příznaku. Obdobným způsobem jsou počítány i ostatní globální příznaky.

Každý vokál je popsán celkem 90 řečovými příznaky (lokálními a vysokoúrovňovými). Počet příznaků pro každou řečovou úlohu databáze PARCZ je 455 (5 vokálů v úloze odpovídá 450 příznakům, navíc je obsaženo 5 globálních). Každého řečníka tedy celkově popisuje 1820 řečových příznaků. Všechny vypočítané příznaky jsou umístěny v matici příznaků, jejíž část je zobrazena v tab. 4.3. Sloupce odpovídají jednotlivým řečníkům a řádky řečovým příznakům. Jejich popis má tvar: "*kód úlohy*"-"*vokál*"-"*lokální příznak(vysokoúrovňový příznak)*" resp. "*kód úlohy*"-"*globální příznak*" pro globální příznaky.

Tab. 4.3: Ukázka části matice příznaků

	K1003	K1004	K1005	K1006	K1007	...
7.1-1-a_ $F_0$ (mean)	203,28	197,65	203,24	193,75	210,33	
7.1-1-a_ $F_0$ (std)	16,79	18,63	11,23	21,57		
7.1-1-a_ $F_0$ (1p)	192,06	174,31	188,86			
7.1-1-a_ $F_0$ (99p)	259,90	233,93				
7.1-1-a_ $F_0$ (ir)	67,84					
⋮						

Tab. 4.4: Přehled implementovaných lokálních příznaků

Zkratka	Typ	Popis
$F_0$	vektor	kmitočet základního tónu
$F_0VR$	skalár	rozdíl maximálních a minimálních hodnot $F_0$
$relF_0SD$	skalár	relativní směrodatná odchylka $F_0$
$relF_0VR$	skalár	relativní kolísání
jitter (rel)	skalár	relativní jitter v %
jitter (abs)	skalár	absolutní jitter
jitter (RAP)	skalár	Relative Average Perturbation
jitter (PPQ5)	skalár	five-point Pitch Perturbation Quotient
shimmer (rel)	skalár	relativní shimmer v %
shimmer (dB)	skalár	shimmer v dB
shimmer (APQ3)	skalár	three-point Amplitude Perturbation Quotient
shimmer (APQ5)	skalár	five-point Amplitude Perturbation Quotient
shimmer (APQ11)	skalár	11-point Amplitude Perturbation Quotient
$\Delta A$	skalár	rozdíl amplitud hlasivkových pulsů
$E$	vektor	krátkodobá energie signálu
$\Delta E$	skalár	rozdíl maximální a minimální hodnoty $E$
TKEO	vektor	Teagrův-Kaiserův energetický operátor
ZCR	vektor	počet průchodů signálu nulovou úrovní
SF	vektor	spektrální tok
SDBM	vektor	změna spektra modulů
SDBP	vektor	změna spektra fází
CPP	vektor	Cepstral Peak Prominence
PPE	skalár	Pitch Period Entropy
HNH	vektor	Harmonic-to-Noise Ratio
NHR	vektor	Noise-to-Harmonic Ratio
RPDE	skalár	Recurrence Probability Density Entropy
DFA	skalár	Detrended Fluctuation Analysis
SE	skalár	Shannonova entropie
RE	skalár	Rényiho entropie
FD	skalár	fraktální dimenze
$F_1$	vektor	1. formantový kmitočet
$F_2$	vektor	2. formantový kmitočet
$FB_1$	vektor	šířka pásma 1. formantu
$FB_2$	vektor	šířka pásma 2. formantu

Tab. 4.5: Přehled implementovaných vysokoúrovňových příznaků

Zkratka	Typ	Popis
mean	skalár	střední hodnota
std	skalár	směrodatná odchylka
1p	skalár	1. percentil
99p	skalár	99. percentil
ir	skalár	absolutní mezipercentilový rozsah mezi 1. a 99. percentilem

Tab. 4.6: Přehled implementovaných globálních příznaků

Zkratka	Typ	Popis
FR	skalár	poměr formantů $F_{2i}/F_{2u}$
FCR	skalár	centralizační poměr formantů
VAI	skalár	artikulační index samohlásek
VSA	skalár	oblast hlasového traktu
lnVSA	skalár	logaritmická oblast hlasového traktu

## 4.3 Statistická analýza

Další částí systému je statistická analýza. Cílem je analyzovat řečové příznaky z hlediska jejich významnosti pro úlohu detekce hypokinetické dysartrie. Mezi neparametrické metody patří Spearmanův test pořadové korelace a Mann-Whitneyho U test. Používají se v případech, kdy nelze předpokládat, že veličiny  $X$ ,  $Y$  mají lineární vztah nebo normální rozdělení. Jejich nevýhodou je slabší síla testu tzn. zamítnutí nepravdivé hypotézy s menší pravděpodobností než parametrickými testy. Dále je implementován výpočet vzájemné informace.

### 4.3.1 Spearmanův test pořadové korelace

Jedná se o neparametrickou metodu, která je založena na výpočtu pořadí hodnot sledovaných veličin. Spearmanův koeficient pořadové korelace je definován jako

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (R_i - Q_i)^2, \quad (4.2)$$

kde  $R_i$  označuje pořadí náhodné veličiny  $X_i$ ,  $Q_i$  pořadí náhodné veličiny  $Y_i$  a  $i = 1, 2, \dots, n$ . Parametr  $n$  určuje počet korelačních dvojic [6]. Výše definovaný vzorec je považován za odhad teoretické hodnoty Spearmanova koeficientu pořadové korelace, která se značí  $\rho_{sp}$ . Pro koeficient platí  $-1 \leq r_s \leq 1$ . Blíže 1 je silnější přímá pořadová závislost mezi veličinami  $X$  a  $Y$ , blíže  $-1$  je silnější nepřímá závislost. Mezi další jeho

vlastnosti patří symetričnost a rezistence vůči odlehlým hodnotám [6]. V prostředí Matlab jeho výpočet realizuje vestavěná funkce `corr` ve tvaru

$$[\text{rho } \text{pval}] = \text{corr}(X, Y, \text{'type'}, \text{'Spearman'}),$$

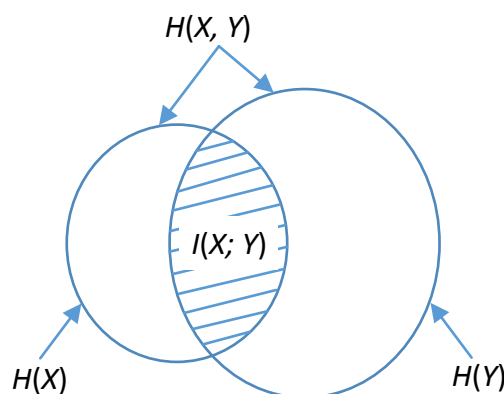
kde  $X$  je vektor příznaků. Vektor  $Y$  hodnotami 0 nebo 1 popisuje, zda daný příznak vektoru  $X$  odpovídá zdravému kontrolnímu řečníkovi (hodnota 0) nebo pacientovi s PN (hodnota 1). Vektor  $Y$  se také nazývá *label* (popisný) vektor. Výstupem funkce je Spearmanův korelační koeficient  $\text{rho}$  a dále hodnota  $\text{pval}$ . Pokud je hodnota  $\text{pval}$  menší než námi stanovená hladina významnosti  $\alpha = 0,05$ , tak je zamítnuta nulová hypotéza  $H_0$ . Ta předpokládá nulový rozdíl mezi testovanými soubory dat. Hladina významnosti definuje pravděpodobnost, že zamítneme  $H_0$ , i když je hypotéza správná [6].

### 4.3.2 Vzájemná informace

Udává velikost informace, kterou náhodná veličina  $X$  obsahuje o druhé náhodné veličině  $Y$ . Veličiny resp. vektory  $X$  a  $Y$  jsou definovány stejným způsobem jako u výpočtu Spearmanova korelačního koeficientu. Pokud je vzájemná informace nulová, tak jsou veličiny statisticky nezávislé. Lze ji vyjádřit jako

$$I(X; Y) = H(X) + H(Y) - H(X, Y), \quad (4.3)$$

kde  $H(X)$  je entropie náhodné veličiny  $X$ , podobně  $H(Y)$  je entropie náhodné veličiny  $Y$  a  $H(X, Y)$  je sdružená entropie  $X$  a  $Y$  [7]. Vztah mezi popsányými veličinami znázorňuje Vennův diagram na obr. 4.3.



Obr. 4.3: Vztah mezi entropií a vzájemnou informací

### 4.3.3 Mann-Whitneyho U test

Další použitou metodou statistického zpracování je Mann-Whitneyho U test. Stejně jako Spearmanův test pořadové korelace patří do skupiny neparametrických testů tzn. je založený na pořadí hodnot. Používá se pro hodnocení nepárových pokusů. Srovnává mediány dvou různých výběrových souborů, které nemusí být početně vyrovnané [6]. Testované hypotézy jsou následující:

- nulová hypotéza  $H_0$  – mediány obou souborů jsou shodné,
- alternativní hypotéza  $H_1$  – mediány jsou rozdílné.

V Matlabu je test realizován vestavěnou funkcí ve tvaru  $p = \text{ranksum}(X, Y)$ , kde vektor  $X$  obsahuje hodnoty příznaků pro kontrolní řečníky a vektor  $Y$  obsahuje hodnoty příznaků pacientů s PN. Pokud při stanovení hladiny významosti  $\alpha = 0,05$  je výstup funkce  $p_{MW} < \alpha$ , zamítáme nulovou hypotézu.

## 4.4 Výběr příznaků

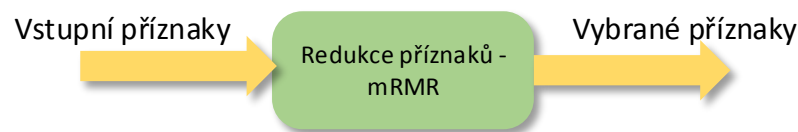
Velké množství příznaků se může stát problémem pro systémy strojového učení. Může vést k větší výpočetní náročnosti, menší efektivitě systému, přetrénování klasifikátoru apod. Z množiny příznaků velikosti  $m$  se vybírá podmnožina velikosti  $n$  takovým způsobem, aby vedla k maximalizaci tzv. kriteriální funkce. Snahou je odstranit nepotřebné a vzájemně korelované příznaky. Proces se označuje jako *selektce příznaků*. Dále se rozděluje na tzv. filter a wrapper metody [16].

### 4.4.1 Metody výběru

Filter metody vybírají podmnožinu příznaků zcela nezávisle na použitém klasifikátoru (obr. 4.4). Jejich výhodou je rychlost a univerzálnost, nicméně poskytují horší výsledky. Do této skupiny patří výběr algoritmem mRMR. Analyzuje vzájemnou informaci mezi diskretizovanými příznaky a jejich třídami (*labely* 0 nebo 1) pro maximální relevanci. Zároveň je hodnocena vzájemná informace mezi diskretizovanými příznaky z důvodu minimalizace redundance. Vektory příznaků jsou kvantovány do tří hladin s rozhodovacími úrovněmi  $\text{mean} \pm \text{std}$ . Výstupem algoritmu jsou příznaky seřazené dle jejich významu a následně se vybere prvních  $n$  příznaků. Podrobné informace o mRMR lze nalézt v pramenu [23].

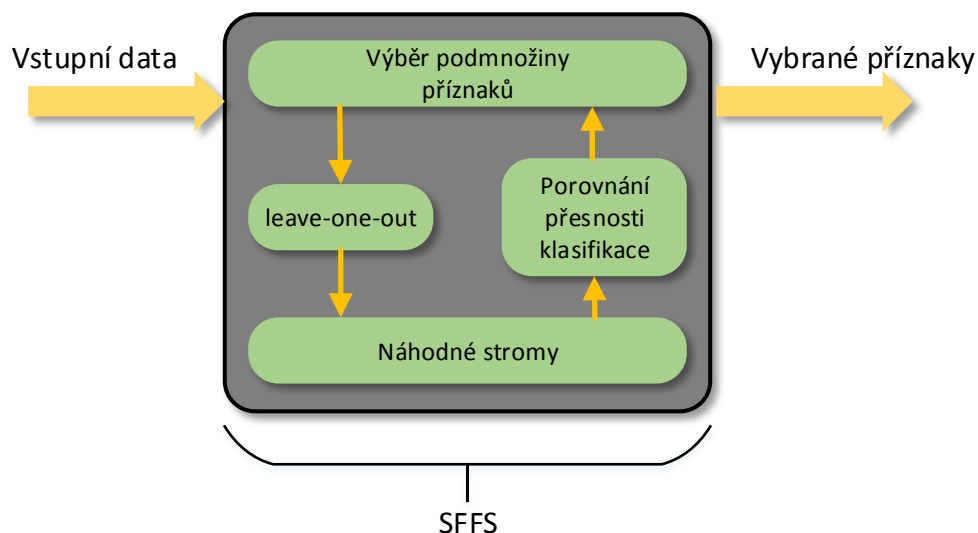
Wrapper metody prohledávají prostor všech možných podmnožin příznaků. Každá uvažovaná podmnožina se otestuje na vstupních datech s použitím klasifikačního algoritmu – viz obr. 4.5. Tento přístup lze chápat jako „černou skříňku“, kdy považujeme za důležitou pouze vybranou skupinu příznaků. Výhodou metody je vyšší klasifikační přesnost. Naopak za nevýhodu lze považovat neuniverzálnost vybraných





Obr. 4.4: Schéma metody filter pro výběr příznaků

příznaků, protože jsou spjaté s konkrétním klasifikátorem. Metoda je také výpočtově náročná – musí se provést trénování a testování klasifikátoru pro všechny možné podmnožiny. Při snaze nalézt optimální řešení zkoušením všech možných kombinací na množině příznaků o velikosti  $m$  by bylo nutné projít  $2^m$  možností. Hledá se tedy suboptimální řešení pomocí heuristických algoritmů.



Obr. 4.5: Schéma metody wrapper pro výběr příznaků

#### 4.4.2 Sekvenční dopředná selekce příznaků

V systému je implementován algoritmus sekvenční dopředné selekce příznaků SFSS (Sequential Forward Floating Search). Jde o plovoucí (*floating*) metodu, u které se dimenzionalita vybrané podmnožiny příznaků nemění monotónně, ale jakoby „plave“ nahoru a dolů. Následující část textu popisuje princip algoritmu včetně zjednodušeného pseudokódu [26].

Prvním krokem algoritmu je inicializace prázdnou množinou  $Y_0$ . V kroku 2 se postupně přidávají příznaky do  $Y_k$  a vypočítá se přesnost klasifikace. (Nejprve se testuje každý příznak samostatně, pak se jedná o dvojice, trojice atd.) Index  $k$  značí velikost množiny  $Y$  tzn. počet příznaků obsažených v množině. Příznak  $x^+$ , který zlepšuje přesnost klasifikace, se přidá do  $Y_k$  (krok 3.1). V kroku 3.2 se nalezne nejméně významný příznak  $x^-$  v množině  $Y_k$ . Pokud je  $x^-$  právě přidáný příznak, ponechá se a pokračuje se krokem 2. Jinak dojde k vyloučení  $x^-$ , které zároveň vede ke zlepšení přesnosti klasifikace. Dále se krokem 3.2 hledá další nejméně významný příznak. Pokud vyloučením  $x^-$  nedojde ke zlepšení, jde se na krok 2 a proces se opakuje. Zastavení běhu algoritmu lze nastavit parametry  $k$  a  $t$ :

1. pokud velikost množiny  $Y_k$  dosáhne počtu příznaků, který jsme předem definovali – např. výběr pěti příznaků stanovením  $k = 5$ ;
2. parametr  $t$  určuje, kolik příznaků je možno zahrnout do  $Y_k$ , aniž by došlo ke zlepšení předchozích výsledků klasifikace. Pokud je např. vybrána množina  $Y_5$  (obsahuje 5 příznaků) a nedojde ke zlepšení výsledků systému ani pro  $Y_{10}$ , algoritmus se ukončí a výsledkem je  $Y_5$ . Parametr  $t$  není z důvodu zjednodušení v pseudokódu uveden.

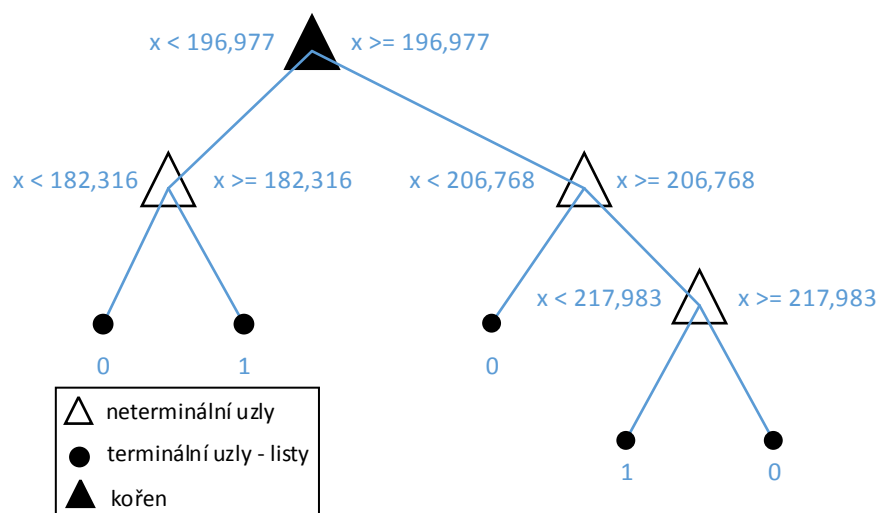
Princip algoritmu SFFS [24, 26]:

1. Inicializace: prázdná množina  
 $Y_0 = \{\}, k = 0$
2. Vyber nejlepší příznak  
 $x^+ = \arg \max_{x^+ \notin Y_k} [J(Y_k + x^+)]$
3. když  $J(Y_k + x^+) > J(Y_k)$ 
  - 3.1. Aktualizuj množinu:  $Y_{k+1} = Y_k + x^+; k = k + 1$
  - 3.2. Odstraň nejhorší příznak:  $x^- = \arg \max_{x^- \in Y_k} [J(Y_k - x^-)]$
  - 3.3. když  $J(Y_k - x^-) > J(Y_k)$ 
    - 3.3.1. Aktualizuj množinu:  $Y_{k+1} = Y_k - x^-; k = k + 1$
    - 3.3.2. Jdi na krok 3.2
  - 3.4. jinak
    - 3.4.1. Jdi na krok 2
4. konec

## 4.5 Klasifikace

### 4.5.1 Klasifikační stromy

Patří mezi neparametrické metody strojového učení. Model lze popsat stromovým grafem, který obsahuje uzly a orientované hrany. U binární stromů se z každého uzlu větví 2 hrany a model klasifikuje vzorky do dvou výstupních tříd. Větvení je v tomto případě založeno na hodnotě jediného prediktoru. Při rozhodování se postupuje od kořene stromu přes neterminální uzly k některému listu. Příklad klasifikačního stromu pro řečový příznak  $F_0$  je na obr. 4.6. Kombinace klasifikačních stromů se označuje jako klasifikační les. Výsledná klasifikační funkce vznikne vhodným zkombinováním klasifikačních funkcí jednotlivých stromů. Pro vytváření lesů se používají metody bagging, boosting a random forests. Další informace lze nalézt v pramenu [15].



Obr. 4.6: Příklad klasifikačního stromu z Matlabu

### 4.5.2 Validace a proces klasifikace

V navrhnutém systému je při klasifikaci použita ve všech případech validace typu vynechání jednoho vzorku (leave-one-out). Její princip je následující: z množiny dat se použije jeden vzorek pro testování a ostatní vzorky pro trénování systému. Tento proces se opakuje tolikrát, kolik je vzorků v celé množině.

Průběh klasifikace implementované v systému lze popsat na následujícím příkladu. Vektor  $X$  obsahuje příznaky a vektor  $Y$  odpovídající labely 0 nebo 1. Validací leave-out-one jsou vektory rozděleny na testovací a trénovací data označená jako

`x_train`, `x_test`, `y_train`, `y_test`. Klasifikace založená na náhodných stromech je provedena funkcemi Matlabu:

```
obj = ClassificationTree.fit(x_train,y_train, 'Prior', 'uniform', ...
                             'SplitCriterion', 'gdi');
y_pred = predict(obj, x_test);
```

První funkce na základě vstupních proměnných natrénuje binární klasifikační strom `obj`. Pro nalezení nejvhodnějšího atributu větvení z kořene stromu je použito kritérium Gini index (nastavením `'SplitCriterion', 'gdi'`). Dalším příkazem se provede predikce na základě natrénovaného modelu a testované hodnoty `x_test`. Výstupem funkce `predict` je hodnota `y_pred`, která odpovídá klasifikované třídě: 0 – negativní výsledek testu (zdravý řečník); 1 – pozitivní výsledek (Parkinsonova nemoc). Poté se porovná skutečný zdravotní stav řečníka `y_test` a jeho predikovaná hodnota `y_pred` – viz tab. 4.7. Výsledek porovnání spadá do některé z následujících kategorií:

- TP (True positive) – nemocný člověk je správně diagnostikován jako nemocný,
- FP (False positive) – zdravý člověk je nesprávně diagnostikován jako nemocný,
- TN (True negative) – zdravý člověk je správně diagnostikován jako zdravý,
- FN (False negative) – nemocný člověk je nesprávně diagnostikován jako zdravý.

Tab. 4.7: Kategorie výsledků testu

		Přítomnost PN	
		Ano (1)	Ne (0)
Výsledek testu	Pozitivní (1)	TP	FP
	Negativní (0)	FN	TN

Po provedené validaci pro všechny vzorky je vypočítána sensitivita (SEN), specifita (SPE) a přesnost klasifikace (ACC) dle vztahů:

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad \text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}. \quad (4.4)$$

Sensitivita je schopnost testu identifikovat skutečně nemocné osoby neboli pravděpodobnost, že test bude pozitivní, když je člověk skutečně nemocný. Specifita je schopnost testu identifikovat osoby bez nemoci tj. pravděpodobnost, že test bude negativní, když člověk není nemocný. Dále je vypočítán parametr nazvaný kompromis mezi sensitivitou a specifikitou TSS [34] (trade-off between sensitivity and specificity)

$$\text{TSS} = 2^{\sin(\frac{\pi \cdot \text{SEN}}{2}) \sin(\frac{\pi \cdot \text{SPE}}{2})}. \quad (4.5)$$

## 5 VÝSLEDKY

Kapitola popisuje dosažené výsledky realizovaného systému. Nejprve jsou řečové příznaky analyzovány jednotlivě. Následně algoritmus SFFS vybírá kombinace příznaků pro dosažení co nejlepšího výsledku klasifikace. Oba způsoby vyhodnocení jsou provedeny pro obě pohlaví dohromady a poté zvlášť pro ženy a muže. Přehled řečových úloh databáze PARCZ popsanych v kap. 4.1 je pro názornost uveden v tab. 5.1.

### 5.1 Individuální analýza příznaků

Jde o první způsob vyhodnocení vhodnosti řečových příznaků z hlediska jejich významu pro klasifikaci hypokinetické dysartrie. Z matice příznaků (tab. 4.3) je vypočítán pro každý řečový příznak Spearmanův korelační koeficient  $\rho_{sp}$  a jeho hladina významnosti  $p_{sp}$ , vzájemná informace MI a hladina významnosti Mann-Whitneyho U testu  $p_{MW}$ . Pro neparametrické testy je stanovena hladina významnosti na  $\alpha = 0,05$ . Dále je provedena validace leave-one-out a klasifikace náhodnými stromy. Její výsledky vyjadřuje přesnost, sensitivita, specificita a také parametr TSS. Pro každou realizaci vokálu je vybrán jeden řečový příznak, který dosáhl nejlepšího kompromisu mezi sensitivitou a specificitou. Tyto výsledky jsou pro obě pohlaví dohromady uvedeny v tab. 5.2, pro ženy v tab. 5.3 a pro muže v tab. 5.4.

Z pohledu statistické analýzy obou pohlaví dohromady má vokál [u] z úlohy 7.1-4 nejvyšší hodnotu Spearmanova korelačního koeficientu a také nejnižší hladiny významností pro oba testy ( $\rho_{sp} = 0,2420$ ,  $p_{sp} = 0,0050$  a  $p_{MW} = 0,0055$ ). Z tab. 5.2 je také patrné, že tento vokál je vyhodnocen jako statisticky významný – společně s vokálem [o] ze stejné úlohy a vokálem [e] z úloh 7.1-1 a 7.1-3. Z pohledu klasifikačních výsledků má nejvyšší hodnotu TSS právě dlouhý vokál [o] z úlohy 7.1-4. Jeho výsledky klasifikace: ACC = 68,42 %, SEN = 67,86 % a SPE = 69,39 %. Mírně vyšší přesnosti klasifikace a významně lepší sensitivity (ACC = 69,92 %, SEN = 76,19 %) dosáhl vokál [i] z úlohy 7.1-3. Nicméně má také znatelně menší specificitu (59,18 %) a z tohoto důvodu také nižší TSS.

V případě žen tento vokál ([i] úloha 7.1-3) dosahuje nejlepších klasifikačních výsledků – ACC = 78,33 %, SEN = 77,78 % a SPE = 79,17 %. Statisticky významných je celkem 5 příznaků. Získané parametry vokálu [u] z úlohy 7.1-2:  $\rho_{sp} = -0,5186$ ,  $p_{sp} = 0,0000$  a  $p_{MW} = 0,0001$ . Také v případě žen byly vokály [o], [u] z úlohy 7.1-4 opět vyhodnoceny jako statisticky významné a navíc [u] dosahuje druhé nejvyšší přesnosti klasifikace a TSS.

Nejvýznamnější je u mužů vokál [u] ze cvičení 7.1-3:  $\rho_{sp} = -0,3261$ ,  $p_{sp} =$

0,0049,  $p_{MW} = 0,0058$ ,  $ACC = 76,71\%$ ,  $SEN = 79,17\%$  a  $SPE = 72,00\%$ . Stejné hodnoty parametru TSS a srovnatelných výsledků dosáhl také vokál [u] z úlohy 7.1-4:  $ACC = 75,34\%$ ,  $SEN = 75,00$  a  $SPE = 76,00\%$ . Mezi statisticky významné patří pouze vokály [o], [u] z úlohy 7.1-3.

Shodným způsobem jsou analyzovány i globální příznaky (tab. 5.5). Dosahují znatelně horších výsledků klasifikace než odpovídající lokální příznaky v jednotlivých řečových úlohách. Žádný globální příznak také není vyhodnocen jako statisticky významný. Z tab. 5.5 vyplývá jasná převaha vybraného poměru formantů  $FR = F_{2i}/F_{2u}$  nad ostatními globálními příznaky. Výsledky klasifikace globálních příznaků dále ukazují, že nejvyšší ACC bylo vždy dosaženo příznaky extrahovanými z úlohy 7.1-4 (pro obě pohlaví, muže i ženy).

Z výše popsáných výsledků vyplývá několik skutečností:

- dlouhé vokály vyslovené s maximální intenzitou (úloha 7.1-3) poskytují nejvyšší ACC – konkrétně [i] (obě pohlaví, ženy) a vokál [u] (muži).
- Velký význam pro analýzu fonace mají dlouhé vokály [o], [u] vyslovené tiše, ale ne šeptem (úloha 7.1-4). Konkrétně vokál [u] je vyhodnocen jako statisticky významný (obě pohlaví, ženy) a také zároveň poskytuje vysokou přesnost klasifikace a vysoké hodnoty TSS (muži, ženy). Tento fakt lze vysvětlit limitní pozicí jazyka při vyslovení vokálu [u] – viz Hellwagův trojúhelník v kap. 2.3. Navíc dlouhá realizace vokálu s minimální intenzitou zvyrazňuje vady hlasu jako tremor, špatné sevření nebo chvění hlasivek apod.
- Globální příznaky dosahují ve všech případech horších výsledků než příznaky lokální odpovídající stejným řečovým úlohám. Dosažené výsledky potvrzují největší význam vokálů z úlohy 7.1-4 z hlediska analýzy fonace.
- Vysoká míra vzájemné informace nemá velkou vypovídající hodnotu o dosažené úspěšnosti klasifikace ani statistické významnosti vokálu.
- Parametr TSS poskytuje vhodný kompromis mezi sensitivitou a specifitou, čímž také zároveň nepřímo vybírá příznak s největší (resp. druhou největší) přesností klasifikace.

Tab. 5.1: Přehled řečových úloh zaměřených na fonaci

kód úlohy	popis úlohy
7.1-1	krátké vyslovení vokálů s přirozenou intenzitou
7.1-2	dlouhé vyslovení vokálů s přirozenou intenzitou
7.1-3	dlouhé vyslovení vokálů s maximální intenzitou
7.1-4	dlouhé vyslovení vokálů s minimální intenzitou, ale ne šeptem

Tab. 5.2: Individuální analýza příznaků – obě pohlaví

Úloha-vokál	Příznak	$\rho_{Sp}$	$p_{Sp}$	MI	$p_{MW}$	ACC (%)	SEN (%)	SPE (%)	TSS
7.1-1-a	TKEO (ir)	0,0524	0,5494	0,6627	0,5489	60,90	63,10	57,14	1,57
7.1-1-e	TKEO (ir)	-0,2209	0,0106	0,7576	0,0112	65,41	67,86	61,22	1,64
7.1-1-i	shimmer (APQ5)	-0,0572	0,5128	0,6768	0,5122	67,67	67,86	67,35	1,70
7.1-1-o	rel $F_0$ SD	-0,0162	0,8528	0,5143	0,8538	66,17	66,67	65,31	1,67
7.1-1-u	ZCR (ir)	-0,1196	0,1701	0,3880	0,1700	63,16	63,10	63,27	1,63
7.1-2-a	SF (mean)	-0,1088	0,2125	0,6805	0,2121	63,91	64,29	63,27	1,64
7.1-2-e	$F_0$ (1p)	-0,0012	0,9889	<b>0,8479</b>	0,9907	66,92	72,62	57,14	1,64
7.1-2-i	HNR (std)	-0,0175	0,8419	0,7764	0,8429	63,91	61,90	67,35	1,65
7.1-2-o	$F_2$ (1p)	-0,0235	0,7879	0,7349	0,7885	64,66	66,67	61,22	1,64
7.1-2-u	ZCR (ir)	-0,0516	0,5553	0,3509	0,5548	66,92	65,48	69,39	1,69
7.1-3-a	shimmer (APQ11)	-0,0106	0,9040	0,7463	0,9053	63,91	66,67	59,18	1,62
7.1-3-e	TKEO (99p)	-0,2257	0,0090	0,5732	0,0096	66,17	70,24	59,18	1,64
7.1-3-i	NHR (mean)	-0,1356	0,1196	0,2839	0,1198	<b>69,92</b>	<b>76,19</b>	<b>59,18</b>	1,68
7.1-3-o	$E$ (ir)	-0,1238	0,1556	0,5350	0,1555	63,91	69,05	55,10	1,59
7.1-3-u	$E$ (99p)	-0,1011	0,2469	0,6647	0,2464	62,41	60,71	65,31	1,62
7.1-4-a	NHR (99p)	-0,1100	0,2074	0,1265	0,2070	63,91	63,10	65,31	1,64
7.1-4-e	NHR (mean)	-0,1482	0,0887	0,2089	0,0891	60,15	60,71	59,18	1,57
7.1-4-i	TKEO (1p)	-0,0276	0,7524	0,7004	0,7529	65,41	64,29	67,35	1,67
7.1-4-o	NHR (std)	-0,2042	0,0184	0,0675	0,0191	<b>68,42</b>	<b>67,86</b>	<b>69,39</b>	<b>1,71</b>
7.1-4-u	HNR (mean)	<b>0,2420</b>	<b>0,0050</b>	0,7162	<b>0,0055</b>	66,17	67,86	63,27	1,66

$\rho_{Sp}$  – Spearmanův korelační koeficient;  $p_{Sp}$  – hladina významnosti (Spearmanův korelační koeficient); MI – vzájemná informace;  $p_{MW}$  – hladina významnosti (Mann-Whitneyův U test); ACC – přesnost klasifikace; SEN – sensitivita; SPE – specifická; TSS – kompromis mezi sensitivitou a specificitou

Tab. 5.3: Individuální analýza příznaků – ženy

Úloha-vokál	Příznak	$\rho_{Sp}$	$p_{Sp}$	MI	$p_{MW}$	ACC (%)	SEN (%)	SPE (%)	TSS
7.1-1-a	$F_0$ (mean)	-0,1198	0,3618	0,8250	0,3613	66,67	63,89	70,83	1,69
7.1-1-e	NHR (99p)	-0,2573	0,0471	0,7332	0,0489	66,67	66,67	66,67	1,68
7.1-1-i	HNR (mean)	0,2554	0,0489	0,7376	0,0507	70,00	72,22	66,67	1,72
7.1-1-o	SDBM (std)	-0,4204	0,0008	0,9376	0,0013	75,00	72,22	79,17	1,81
7.1-1-u	shimmer (rel)	0,1355	0,3018	0,8710	0,3013	71,67	69,44	75,00	1,76
7.1-2-a	HNR (1p)	0,1945	0,1365	0,8917	0,1372	70,00	72,22	66,67	1,72
7.1-2-e	$F_0$ (1p)	-0,0688	0,6017	0,8584	0,6027	75,00	75,00	75,00	1,81
7.1-2-i	$FB_2$ (1p)	0,1414	0,2810	0,8125	0,2806	73,33	75,00	70,83	1,78
7.1-2-o	ZCR (mean)	0,0707	0,5913	0,8917	0,5922	70,00	72,22	66,67	1,72
7.1-2-u	SF (mean)	<b>-0,5186</b>	<b>0,0000</b>	0,8710	<b>0,0001</b>	71,67	75,00	66,67	1,74
7.1-3-a	SDBM (mean)	0,1532	0,2425	0,8376	0,2422	68,33	77,78	54,17	1,63
7.1-3-e	$F_2$ (ir)	0,0000	1,0000	0,9376	1,0000	73,33	77,78	66,67	1,76
7.1-3-i	HNR (99p)	-0,0648	0,6226	0,8917	0,6238	<b>78,33</b>	<b>77,78</b>	<b>79,17</b>	<b>1,85</b>
7.1-3-o	$FB_1$ (ir)	-0,2239	0,0854	0,7502	0,0868	76,67	83,33	66,67	1,79
7.1-3-u	ZCR (99p)	0,0698	0,5961	0,4567	0,5971	66,67	61,11	75,00	1,69
7.1-4-a	NHR (std)	-0,2357	0,0698	0,7900	0,0714	73,33	75,00	70,83	1,78
7.1-4-e	SF (99p)	-0,1964	0,1325	<b>0,9710</b>	0,1333	68,33	66,67	70,83	1,71
7.1-4-i	HNR (std)	0,0079	0,9525	0,8376	0,9579	68,33	66,67	70,83	1,71
7.1-4-o	shimmer (APQ11)	-0,3497	0,0062	0,8376	0,0074	68,33	75,00	58,33	1,66
7.1-4-u	RPDE	-0,3988	0,0016	0,8710	0,0022	76,67	77,78	75,00	1,83

$\rho_{Sp}$  – Spearmanův korelační koeficient;  $p_{Sp}$  – hladina významnosti (Spearmanův korelační koeficient); MI – vzájemná informace;  $p_{MW}$  – hladina významnosti (Mann-Whitneyův U test); ACC – přesnost klasifikace; SEN – sensitivita; SPE – specifická; TSS – kompromis mezi sensitivitou a specifícií



Tab. 5.4: Individuální analýza příznaků – muži

Úloha-vokál	Příznak	$\rho_{Sp}$	$p_{Sp}$	MI	$p_{MW}$	ACC (%)	SEN (%)	SPE (%)	TSS
7.1-1-a	HNR (std)	0,1781	0,1317	<b>0,8620</b>	0,1322	64,38	62,50	68,00	1,66
7.1-1-e	SF (std)	-0,0849	0,4749	0,7798	0,4747	72,60	75,00	68,00	1,75
7.1-1-i	NHR (std)	-0,1699	0,1508	0,4933	0,1511	72,60	75,00	68,00	1,75
7.1-1-o	$F_1$ (1p)	0,0192	0,8720	0,8279	0,8753	72,60	77,08	64,00	1,73
7.1-1-u	HNR (1p)	0,0973	0,4130	0,7421	0,4125	69,86	72,92	64,00	1,70
7.1-2-a	RPDE	-0,2123	0,0713	0,8176	0,0725	71,23	72,92	68,00	1,74
7.1-2-e	jitter (abs)	-0,1370	0,2478	0,6976	0,2474	68,49	66,67	72,00	1,72
7.1-2-i	$FB_2$ (std)	-0,1466	0,2159	0,8072	0,2157	68,49	68,75	68,00	1,71
7.1-2-o	HNR (std)	-0,1151	0,3323	0,7798	0,3317	65,75	62,50	72,00	1,68
7.1-2-u	ZCR (99p)	0,0536	0,6524	0,3175	0,6534	69,86	66,67	76,00	1,75
7.1-3-a	ZCR (mean)	-0,0548	0,6452	0,7969	0,6461	71,23	68,75	76,00	1,77
7.1-3-e	$E$ (mean)	0,0384	0,7473	0,6806	0,7492	73,97	75,00	72,00	1,79
7.1-3-i	$F_1$ (mean)	-0,1411	0,2337	0,6909	0,2334	72,60	72,92	72,00	1,77
7.1-3-o	HNR (ir)	-0,2370	0,0435	0,7591	0,0449	75,34	79,17	68,00	1,78
7.1-3-u	NHR (99p)	<b>-0,3261</b>	<b>0,0049</b>	0,3773	<b>0,0058</b>	<b>76,71</b>	<b>79,17</b>	<b>72,00</b>	<b>1,81</b>
7.1-4-a	SDBM (99p)	-0,1082	0,3621	0,7969	0,3615	69,86	77,08	56,00	1,65
7.1-4-e	NHR (mean)	-0,0616	0,6044	0,3174	0,6049	73,97	72,92	76,00	1,80
7.1-4-i	TKEO (1p)	-0,0041	0,9725	0,7147	0,9768	67,12	68,75	64,00	1,68
7.1-4-o	$F_2$ (1p)	-0,1466	0,2159	0,7147	0,2157	69,86	68,75	72,00	1,74
7.1-4-u	SDBP (mean)	0,2096	0,0751	0,7318	0,0763	<b>75,34</b>	<b>75,00</b>	<b>76,00</b>	<b>1,81</b>

$\rho_{Sp}$  – Spearmanův korelační koeficient;  $p_{Sp}$  – hladina významnosti (Spearmanův korelační koeficient); MI – vzájemná informace;  $p_{MW}$  – hladina významnosti (Mann-Whitneyův U test); ACC – přesnost klasifikace; SEN – sensitivita; SPE – specificita; TSS – kompromis mezi sensitivitou a specificitou

Tab. 5.5: Individuální analýza příznaků – globální příznaky

	Úloha	Příznak	$\rho_{Sp}$	$p_{Sp}$	MI	$p_{MW}$	ACC (%)	SEN (%)	SPE (%)	TSS
Obě pohlaví	7.1-1	FR	-0,0942	0,2809	0,7512	0,2802	56,39	59,52	51,02	1,49
	7.1-2	FR	-0,0097	0,9114	0,7142	0,9127	54,14	55,95	51,02	1,47
	7.1-3	FCR	-0,0426	0,6261	0,7012	0,6259	52,63	57,14	44,90	1,42
	7.1-4	FR	-0,0828	0,3433	0,5128	0,3425	<b>60,15</b>	<b>61,90</b>	<b>57,14</b>	<b>1,56</b>
Ženy	7.1-1	lnVSA	0,0786	0,5507	0,8043	0,5512	53,33	52,78	54,17	1,47
	7.1-2	FR	-0,1277	0,3309	0,8917	0,3304	66,67	63,89	70,83	1,69
	7.1-3	FR	0,1414	0,281	0,7376	0,2806	51,67	58,33	41,67	1,40
	7.1-4	lnVSA	-0,0452	0,7318	0,9376	0,7342	<b>73,33</b>	<b>72,22</b>	<b>75,00</b>	<b>1,79</b>
Muži	7.1-1	FR	-0,1301	0,2724	0,7731	0,2720	54,79	54,17	56,00	1,49
	7.1-2	FR	0,0685	0,5647	0,6976	0,5650	60,27	64,58	52,00	1,54
	7.1-3	FR	0,0329	0,7824	0,6585	0,7847	49,32	52,08	44,00	1,38
	7.1-4	FR	-0,0068	0,9541	0,7524	0,9583	<b>65,75</b>	<b>66,67</b>	<b>64,00</b>	<b>1,66</b>

$\rho_{Sp}$  – Spearmanův korelační koeficient;  $p_{Sp}$  – hladina významnosti (Spearmanův korelační koeficient); MI – vzájemná informace;  $p_{MW}$  – hladina významnosti (Mann-Whitneyův U test); ACC – přesnost klasifikace; SEN – sensitivita; SPE – specifická; TSS – kompromis mezi sensitivitou a specifícií

## 5.2 Kombinace řečových příznaků

Dalším způsobem vyhodnocení funkčnosti systému je výběr řečových příznaků pomocí algoritmu SFFS. V tomto případě se z matice příznaků nejprve odstraní globální příznaky, které nejsou v tomto typu testování zahrnuty z důvodu jejich neuspokojivých výsledků v předchozí úloze. Výsledky systému jsou opět vyhodnoceny pro každé pohlaví jednotlivě i dohromady ve třech možných scénářích:

1. výběr příznaků pro každý vokál samostatně,
2. výběr příznaků pro každou řečovou úlohu,
3. výběr příznaků ze všech 20 realizací vokálů.

SFFS používá kombinaci klasifikátoru náhodných stromů s validací leave-one-out. Ve 3. scénáři byly příznaky nejprve redukovány pomocí mRMR na 500 příznaků (z původních 1 800). Poté byl proveden jejich výběr algoritmem SFFS stejným způsobem jako u prvních dvou scénářů. Redukce příznaků pomocí mRMR je provedena z důvodu menší časové výpočetní náročnosti při následném nasazení SFFS. Při nastavení algoritmu SFFS (viz kap. 4.4.2) parametrem  $k$  nebyl omezen maximální počet vybraných příznaků a  $t = 5$ . Výsledky všech scénářů jsou v tab. 5.6, 5.7 a 5.8 vyjádřeny přesností klasifikace, sensitivitou, specificitou, parametrem TSS a dále je uveden počet vybraných příznaků. Pro nejlepší dosažené výsledky klasifikace každého scénáře (zvýrazněné v tabulkách tučně) jsou v tab. 5.9 uvedeny konkrétní příznaky vybrané algoritmem SFFS. Je vždy označen název příznaku, případně mu předchází označení vokálu resp. úlohy, ze které byl extrahován (pokud tato informace není zřejmá ze sloupce tabulky Úloha-vokál).

V prvním scénáři pro obě pohlaví dohromady má nejvyšší přesnosti klasifikace krátký vokál [a] vyslovený s přirozenou intenzitou: ACC = 84,21 %, SEN = 91,67 %, SPE = 71,43 % s výběrem 6 příznaků. Druhé nejvyšší ACC a zároveň lepšího parametru TSS dosáhl vokál [u] z úlohy 7.1-4. V případě nejlepšího výsledku druhého scénáře nedošlo u úlohy 7.1-3 ke zlepšení žádného parametru ve srovnání s prvním scénářem (ACC = 81,95%, SEN = 89,29%, SPE = 69,39%). Úloha 7.1-4 má srovnatelnou ACC s úlohou 7.1-3 a také lepší TSS. Výběrem příznaků ze všech realizací vokálů v rámci třetího scénáře bylo dosaženo nejvyšší přesnosti klasifikace – ACC = 89,47%, SEN = 91,67 % a SPE = 85,71 %. Došlo k procentuálnímu nárůstu u všech parametrů, včetně TSS. Ze 14 použitých příznaků je celkem 5 vybráno z úlohy 7.1-4 – viz tab. 5.9.

U žen v tab. 5.7 bylo dosaženo stejně vysoké úspěšnosti klasifikace v prvním scénáři pro tři realizace vokálů (ACC = 90,00 %, SEN = 91,67 %, SPE = 87,50 %). Liší se pouze počtem vybraných příznaků, kterých je nejméně u vokálu [e] z úlohy 7.1-1. Podobně jako v případě obou pohlaví se zde vyskytuje realizace vokálu [u] z úlohy 7.1-4. U druhého scénáře došlo u úlohy 7.1-4 k dalšími zlepšení výsledků, dokonce

i se zachováním výběru pouhých tří příznaků:  $ACC = SEN = SPE = 91,67\%$ . Je to také případ nejlepšího dosaženého kompromisu mezi sensitivitou a specificitou. Stejně jako v případě obou pohlaví dohromady, i u žen došlo v posledním scénáři k dalšímu nárůstu přesnosti klasifikace a především sensitivity:  $ACC = 93,33\%$ ,  $SEN = 97,22\%$  a  $SPE = 87,50\%$ .

U mužů v prvním scénáři mají dvě realizace vokálu [u] z úloh 7.1-1 a 7.1-2 stejné výsledky:  $ACC = 89,04\%$ ,  $SEN = 93,75\%$  a  $SPE = 80,00\%$ . Stejnou  $ACC$ , ale lepší  $TSS$  dosahuje pro muže třetí testovaný scénář:  $ACC = 89,04\%$ ,  $SEN = 89,58\%$  a  $SPE = 88,00\%$ . Přesto se nejedná o nejlepší výsledek – toho bylo dosaženo ve druhém scénáři řečovými příznaky z úlohy 7.1-2, se kterými  $ACC = 91,78\%$ ,  $SEN = 93,75\%$  a  $SPE = 88,00\%$ . Výsledky mužů se liší v jednom ohledu – úloha 7.1-4 u mužů poskytuje nejhorší výsledky v rámci druhého scénáře, zatímco u žen a obou pohlaví dohromady vždy patří tato úloha k nejvýznamnějším. Obdobná vlastnost je také jasně patrná u vokálu [u] ze stejné úlohy.

Výsledky systému lze shrnout do následujících bodů:

- kombinování příznaků přirozeně vede k významně lepším výsledkům klasifikace. Realizovaný systém prokazuje dostatečnou úroveň přesnosti klasifikace, sensitivity i specificity a to nejen pro jednotlivá pohlaví zvlášť, ale i dohromady.
- Systém je schopen rozpoznat řeč pacientů s PN od zdravých kontrolních řečníků s vysokou pravděpodobností. Nejlepších  $ACC$  dosáhl systém pro ženy, následují muži a poté obě pohlaví dohromady.
- Nejvyšší dosažené výsledky: obě pohlaví –  $ACC = 89,47\%$ ,  $SEN = 91,67\%$ ,  $SPE = 85,71\%$ ; ženy –  $ACC = 93,33\%$ ,  $SEN = 97,22\%$ ,  $SPE = 87,50\%$ ; muži –  $ACC = 91,78\%$ ,  $SEN = 93,75\%$ ,  $SPE = 88,00\%$ .
- Výsledky klasifikace jsou lepší než srovnatelné práce ostatních autorů – pokud uvažujeme velikost řečového korpusu, která je srovnatelná s použitou databází PARCZ (133 řečníků). Např. autoři Belalcazar et al. [3] dosáhli pouze  $ACC = 66,50\%$  a Hazan et al. [10]  $ACC = 84,69\%$  (počet řečníků v databázi 100 resp. 98) – viz tab. 3.1.
- Vokál [u] z úlohy 7.1-4 stejně jako úloha samotná patří ve svých scénářích mezi nejvýznamnější pro ženy a obě pohlaví dohromady. Výsledky ukazují, že tato realizace vokálu má pro analýzu fonace velký přínos. Toto zjištění odpovídá poznatkům z individuální analýzy příznaků.
- Nízké počty použitých příznaků (maximálně 14) jsou způsobeny jejich efektivním výběrem algoritmem SFFS.

Tab. 5.6: Výsledky klasifikace – obě pohlaví

Úloha-vokál	ACC (%)	SEN (%)	SPE (%)	TSS	No.
7.1-1-a	<b>84,21</b>	<b>91,67</b>	<b>71,43</b>	1,86	6
7.1-1-e	77,44	78,57	75,51	1,83	10
7.1-1-i	75,94	80,95	67,35	1,78	11
7.1-1-o	71,43	82,14	53,06	1,64	3
7.1-1-u	79,70	85,71	69,39	1,82	6
7.1-2-a	74,44	77,38	69,39	1,78	5
7.1-2-e	74,44	79,76	65,31	1,76	6
7.1-2-i	75,19	83,33	61,22	1,73	13
7.1-2-o	72,18	82,14	55,10	1,66	6
7.1-2-u	81,95	85,71	75,51	1,87	8
7.1-3-a	79,70	82,14	75,51	1,85	8
7.1-3-e	73,68	82,14	59,18	1,71	9
7.1-3-i	77,44	84,52	65,31	1,78	3
7.1-3-o	81,95	88,10	71,43	1,85	7
7.1-3-u	73,68	82,14	59,18	1,71	6
7.1-4-a	74,44	82,14	61,22	1,73	3
7.1-4-e	78,20	77,38	79,59	1,85	11
7.1-4-i	81,20	83,33	77,55	1,87	10
7.1-4-o	74,44	75,00	73,47	1,80	6
7.1-4-u	82,71	83,33	81,63	<b>1,90</b>	5
7.1-1	80,45	82,14	77,55	1,87	11
7.1-2	80,45	84,52	73,47	1,85	7
7.1-3	<b>81,95</b>	<b>89,29</b>	<b>69,39</b>	1,83	5
7.1-4	81,20	82,14	79,59	<b>1,88</b>	4
vše	<b>89,47</b>	<b>91,67</b>	<b>85,71</b>	<b>1,95</b>	14

ACC – přesnost klasifikace; SEN – sensitivita; SPE – specificita; TSS – kompromis mezi sensitivitou a specificitou; No. – počet vybraných příznaků

Tab. 5.7: Výsledky klasifikace – ženy

Úloha-vokál	ACC (%)	SEN (%)	SPE (%)	TSS	No.
7.1-1-a	78,33	83,33	70,83	1,82	4
7.1-1-e	<b>90,00</b>	<b>91,67</b>	<b>87,50</b>	<b>1,96</b>	3
7.1-1-i	<b>90,00</b>	<b>91,67</b>	<b>87,50</b>	<b>1,96</b>	5
7.1-1-o	86,67	86,11	87,50	1,94	4
7.1-1-u	83,33	86,11	79,17	1,90	5
7.1-2-a	85,00	80,56	91,67	1,93	4
7.1-2-e	86,67	88,89	83,33	1,93	7
7.1-2-i	76,67	80,56	70,83	1,81	2
7.1-2-o	81,67	86,11	75,00	1,87	2
7.1-2-u	88,33	97,22	75,00	1,90	4
7.1-3-a	83,33	86,11	79,17	1,90	4
7.1-3-e	80,00	80,56	79,17	1,87	5
7.1-3-i	78,33	77,78	79,17	1,85	1
7.1-3-o	78,33	83,33	70,83	1,82	1
7.1-3-u	85,00	86,11	83,33	1,92	7
7.1-4-a	78,33	80,56	75,00	1,84	4
7.1-4-e	83,33	91,67	70,83	1,85	4
7.1-4-i	81,67	88,89	70,83	1,84	3
7.1-4-o	83,33	88,89	75,00	1,88	5
7.1-4-u	<b>90,00</b>	<b>91,67</b>	<b>87,50</b>	<b>1,96</b>	4
7.1-1	88,33	94,44	79,17	1,92	4
7.1-2	86,67	94,44	75,00	1,89	8
7.1-3	83,33	86,11	79,17	1,90	2
7.1-4	<b>91,67</b>	<b>91,67</b>	<b>91,67</b>	<b>1,98</b>	3
vše	<b>93,33</b>	<b>97,22</b>	<b>87,50</b>	<b>1,97</b>	3

ACC – přesnost klasifikace; SEN – sensitivita; SPE – specificita; TSS – kompromis mezi sensitivitou a specificitou; No. – počet vybraných příznaků

Tab. 5.8: Výsledky klasifikace – muži

Úloha-vokál	ACC (%)	SEN (%)	SPE (%)	TSS	No.
7.1-1-a	83,56	85,42	80,00	1,90	7
7.1-1-e	84,93	89,58	76,00	1,89	5
7.1-1-i	83,56	89,58	72,00	1,86	5
7.1-1-o	78,08	87,50	60,00	1,73	3
7.1-1-u	<b>89,04</b>	<b>93,75</b>	<b>80,00</b>	<b>1,93</b>	5
7.1-2-a	86,30	87,50	84,00	1,93	8
7.1-2-e	78,08	81,25	72,00	1,82	8
7.1-2-i	75,34	75,00	76,00	1,81	3
7.1-2-o	78,08	83,33	68,00	1,80	8
7.1-2-u	<b>89,04</b>	<b>93,75</b>	<b>80,00</b>	<b>1,93</b>	4
7.1-3-a	79,45	85,42	68,00	1,81	4
7.1-3-e	80,82	89,58	64,00	1,78	6
7.1-3-i	83,56	85,42	80,00	1,90	5
7.1-3-o	87,67	91,67	80,00	1,92	6
7.1-3-u	82,19	89,58	68,00	1,82	2
7.1-4-a	80,82	87,50	68,00	1,81	5
7.1-4-e	80,82	85,42	72,00	1,84	2
7.1-4-i	84,93	85,42	84,00	1,92	4
7.1-4-o	82,19	83,33	80,00	1,89	8
7.1-4-u	84,93	89,58	76,00	1,89	5
7.1-1	86,30	91,67	76,00	1,89	5
7.1-2	<b>91,78</b>	<b>93,75</b>	<b>88,00</b>	<b>1,97</b>	6
7.1-3	89,04	93,75	80,00	1,93	5
7.1-4	83,56	93,75	64,00	1,79	7
vše	<b>89,04</b>	<b>89,58</b>	<b>88,00</b>	<b>1,96</b>	5

ACC – přesnost klasifikace; SEN – sensitivita; SPE – specificita; TSS – kompromis mezi sensitivitou a specificitou; No. – počet vybraných příznaků

Tab. 5.9: Přehled vybraných řečových příznaků algoritmem SFFS

	Úloha-vokál	Vybrané příznaky				
Obě pohlaví	7.1-1-a	$FB_1(\text{ir})$ HNR(99p)	SDBM(std)	NHR(std)	ZCR(1p)	$FB_1(1p)$
	7.1-4-u	HNR(mean)	$FB_2(99p)$	$FB_2(\text{ir})$	FD	SDBP(99p)
	7.1-3	i_NHR(mean)	i_CPP(ir)	u_ $FB_1(99p)$	u_NHR(std)	i_ $E(\text{mean})$
	7.1-4	o_NHR(std)	e_HNR(1p)	o_ $\Delta A$	a_ $F_2(\text{mean})$	
	vše	7.1-3-i_rel $F_0$ SD 7.1-4-u_ $F_2(\text{std})$ 7.1-4-u_ZCR(1p)	7.1-4-i_shimmer(APQ3) 7.1-1-i_ $F_2(1p)$ 7.1-2-i_SF(99p)	7.1-1-o_ $F_1(99p)$ 7.1-4-i_ $F_1(\text{ir})$ 7.1-3-i_jitter(RAP)	7.1-2-o_ZCR(1p) 7.1-4-e_shimmer(rel) 7.1-3-i_ $FB_2(\text{std})$	7.1-2-a_ $F_1(\text{ir})$ 7.1-2-i_SF(std)
Ženy	7.1-1-e	NHR(99p)	SF(1p)	CPP(std)		
	7.1-1-i	HNR(mean)	rel $F_0$ VR	$F_1(\text{ir})$	ZCR(mean)	$F_2(\text{mean})$
	7.1-4-u	RPDE	$F_1(\text{std})$	$F_0(1p)$	$E(1p)$	
	7.1-4	u_RPDE	o_ $F_1(1p)$	e_RE		
	vše	7.1-4-u_RPDE	7.1-3-u_ $F_2(\text{mean})$	7.1-1-i_SDBM(std)		
Muži	7.1-1-u	$F_0(99p)$	NHR(99p)	$F_0(1p)$	ZCR(99p)	$E(1p)$
	7.1-2-u	HNR(99p)	NHR(mean)	NHR(std)	shimmer(rel)	
	7.1-2	a_RPDE i_ $E(99p)$	o_NHR(1p)	o_ $F_0(\text{std})$	e_SDBM(std)	a_ $F_0(99p)$
	vše	7.1-3-o_HNR(ir)	7.1-4-i_shimmer(rel)	7.1-3-u_SF(std)	7.1-4-o_PPE	7.1-3-e_ $F_1(99p)$



## 6 ZÁVĚR

V této diplomové práci byla nejprve popsána hypokinetická dysartrie jako souhrnné označení nejčastější řečové poruchy u pacientů s Parkinsonovou nemocí. Hypokinetická dysartrie ovlivňuje složky vytvářející řeč – oblast fonace, artikulace, prozodie, plynulosti řeči a faciokineze.

Pomocí řečových příznaků lze patologickou řeč popsat a následně vhodným systémem diagnostikovat zdravotní stav pacienta. Tento neinvazivní diagnostický přístup pomocí analýzy řečových signálů je velmi perspektivní nejen z důvodu rostoucího počtu lidí s neurodegenerativními onemocněními. Je výhodný především díky své objektivitě, nezávislosti na lidském faktoru, ale také rychlosti a nízkých nákladech. Funkčnost a spolehlivost systému na druhou stranu závisí na implementovaných metodách, řečových příznacích a především také na použitém řečovém korpusu, jeho velikosti a kvalitě.

V kapitole 3 jsou popsány různé přístupy ostatních autorů k návrhu takového systému. Nicméně jejich výsledky převážně nelze považovat za dostatečně robustní. Především často pracují s databázemi s malým počtem řečníků nebo se zaměřují na analýzu malé skupiny vokálů nebo řečových příznaků.

Návrhu robustního systému pro analýzu českých vokálů se věnuje kap. 4. První výhodou je použití databáze PARCZ, která je nahrávána ve Fakultní nemocnici u sv. Anny v Brně. Obsahuje celkem 133 řečníků – 84 pacientů s Parkinsonovou nemocí a odpovídajících 49 kontrolních řečníků. Její část s nahranými vokály obsahuje pět českých vokálů ve čtyřech různých řečových úlohách. Použití různých realizací vokálů (dlouhé, krátké, s minimální intenzitou apod.) značně rozšiřuje možnosti analýzy fonace. Celkem 90 řečových příznaků je systémem extrahováno z každé realizace vokálu. Jsou zastoupeny příznaky pro popis fonace, kvality hlasu, odhad úrovně šumu, kmitání hlasivek a také některé nelineární dynamické parametry.

Výsledky systému vytvořeného v prostředí Matlab jsou v kap. 5 vyhodnoceny dvěma způsoby. Prvním je individuální analýza, ve které jsou vokály nejprve podrobeny statistickým testům (Spearmanův korelační test, vzájemná informace a Mann-Whitneyho U test) a následně klasifikaci náhodnými stromy s validací typu leave-one-out. V této analýze dosáhly nejvyšších přesností klasifikace vokály vyslovené dlouze s maximální intenzitou: obě pohlaví (vokál [i]) – ACC = 69,92 %, SEN = 76,19 %, SPE = 59,18 %; ženy (vokál [i]) – ACC = 78,33 %, SEN = 77,78 %, SPE = 79,17 %; muži (vokál [u]) – ACC = 76,71 %, SEN = 79,17 % a SPE = 72,00 %. Zcela srovnatelné výsledky ACC a často lepších hodnot SEN a SPE dosáhly vokály ([o], [u]) vyslovené tiše s minimální intenzitou, ale ne šeptem. Především vokál [u] v tomto typu úlohy pro jednotlivá pohlaví dosahuje jedny z nejlepších kompromisů mezi sensitivitou a specifitou. Při artikulaci [u] se jazyk nachází v limitní pozici

a tím pádem se nejvíce projevují a zvýrazňují hlasové vady jako tremor, nesprávná činnost hlasivkových svalů apod. Výsledky práce ukázaly, že v těchto realizacích (maximální a minimální intenzita dlouhých vokálů) jsou projevy hypokinetické dysartrie nejvíce patrné a mají tedy největší význam z hlediska klasifikace Parkinsonovy nemoci. Druhým způsobem vyhodnocení je výběr příznaků algoritmem SFFS ve třech možných scénářích: 1) pro každý vokál; 2) pro každou řečovou úlohu; 3) výběr ze všech realizací vokálů. Tímto způsobem došlo k výraznému zlepšení výsledků: nejvyšší dosažené přesnosti klasifikace dosáhly samostatně ženy ve 3. scénáři ( $ACC = 93,33\%$ ,  $SEN = 97,22\%$ ,  $SPE = 87,50\%$ ), následně muži ve 2. scénáři ( $ACC = 91,78\%$ ,  $SEN = 93,75\%$ ,  $SPE = 88,00\%$ ) a poté obě pohlaví dohromady ve 3. scénáři ( $ACC = 89,47\%$ ,  $SEN = 91,67\%$ ,  $SPE = 85,71\%$ ). I v tomto způsobu vyhodnocení patřila realizace vokálu [u] k těm nejvýznamnějším, podobně jako v předchozím případě.

Navržený systém dosahuje výsledků, které zaručují nejen vysokou přesnost klasifikace, ale i sensitivity a specificity. Při porovnání s ostatními autory systém dokonce vykazuje lepší výsledky klasifikace, pokud uvažujeme srovnatelnou velikost použitých řečových korpusů.

Funkčnost systému je možné rozšířit a zdokonalit implementací dalších typů řečových příznaků nebo testováním odlišných řečových korpusů (např. v jiných jazycích). Další možností je rozšíření o určení stupně onemocnění regresními stromy pomocí odhadu hodnot různých škál jako je UPDRS (unifikovaná škála pro hodnocení Parkinsonovy nemoci) nebo MMSE (Mini-Mental State Examination).

## LITERATURA

- [1] ALONSO, J. B., DE LEON, J., ALONSO, I., FERRER, M. A. Automatic Detection of Pathologies in the Voice by HOS Based Parameters. *EURASIP J Adv Sig Pr*, ročník 2001, č. 4, 2001, s. 275–284.
- [2] ARROYAVE, J. R. O., ARIAS-LONDONO, J. D., BONILLA, J. F. V., NOTH, E. Analysis of Speech from People with Parkinson's Disease through Nonlinear Dynamics. In *Advances in Nonlinear Speech Processing, Lecture Notes in Computer Science*, editace T. Drugman, T. Dutoit, Springer Berlin Heidelberg, 2013, s. 112–119.
- [3] BELALCAZAR-BOLANOS, E., OROZCO-ARROYAVE, J., ARIAS-LONDONO, J., VARGAS-BONILLA, J., NOTH, E. Automatic detection of Parkinson's disease using noise measures of speech. In *Image, Signal Processing, and Artificial Vision (STSIVA), 2013 XVIII Symposium of*, 2013, s. 1–5.
- [4] BOERSMA P., WEENINK D. Praat: doing phonetics by computer. [Online]. [cit. 2015-05-08]. Dostupné z <http://www.fon.hum.uva.nl/praat/>
- [5] BROOKES, M. VOICEBOX: Speech Processing Toolbox for MATLAB. [Online]. [cit. 2015-05-08]. Dostupné z <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>
- [6] BUDÍKOVÁ, M., KRÁLOVÁ M., MAROŠ, M. *Průvodce základními statistickými metodami*. 1. vyd. Praha: Grada, 2010, 272 s. Expert (Grada). ISBN 978-80-247-3243-5.
- [7] COVER, T., THOMAS, J. *Elements of information theory*. New York: John Wiley & Sons, c1991, xxii, 542 s. Wiley series in telecommunications. ISBN 0471062596.
- [8] DELIYSKI, D. D. Acoustic Model and Evaluation of Pathological Voice Production. In *3rd Conference on Speech Communication and Technology EUROSPEECH'93*, 1993, s. 1969–1972.
- [9] FARRÚS, M., HERNANDO, J., EJARQUE, P. Jitter and Shimmer Measurements for Speaker Recognition. In *INTERSPEECH*. 2007, str. 778–781.
- [10] HAZAN, H., HILU, D., MANEVITZ, L., RAMIG, L., SAPIR, S. Early diagnosis of Parkinson's disease via machine learning on speech data. In *Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, 2012, s. 1–4.

- [11] HENRIQUEZ, P., ALONSO, J.B., FERRER, M.A., TRAVIESO, C.M., GODINO-LLORENTE, J.I., DIAZ-DE-MARIA, F. Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics. *IEEE T Audio Speech*, ročník 17, č. 6, 2009, s. 1186–1195.
- [12] HILLENBRAND, J. Cepstral peak prominence. *Western Michigan University*. [online]. 2002 [cit. 2015-05-08]. Dostupné z <http://homepages.wmich.edu/~hillenbr/>
- [13] HILLENBRAND, J., HOUDE R. A. Acoustic Correlates of Breathy Vocal Quality: Dysphonic Voices and Continuous Speech. *J Speech Hear Res*, ročník 39, č. 2, 1996, s. 311–321.
- [14] KAKINO, N., FUKUMORI, T., NAKAYAMA, M., NISHIURA T. Experimental study of shout detection with the rahmonic structure. In *Proceedings of Meetings on Acoustics*, ICA 2013 Montreal, 2–7 June, 2013.
- [15] KLASCHKA, J., KOTRČ, E. Klasifikační a regresní lesy. *ROBUST 2004. Sborník prací 13. letní školy JČMF*, 2004.
- [16] KOHAVI, R., JOHN, G.H. Wrappers for feature subset selection. *Artificial Intelligence*, ročník 97, č. 1–2, (December 1997), s. 273–324.
- [17] KOŠTÁLOVÁ, M., MRAČKOVÁ, M., MAREČEK, R., BERÁNKOVÁ, D., ELIÁŠOVÁ, I., JANOUŠOVÁ, E., ROUBÍČKOVÁ, J., BEDNAŘÍK, J., REKTOROVÁ, I. Test 3F Dysartrický profil – normativní hodnoty řeči v češtině. *Cesk Slov Neurol N*, 2013, ročník 76/109, č. 5, s. 614—618.
- [18] LITTLE, M., McSHARRY, P., HUNTER, E., SPIELMAN, J., RAMIG, L. Suitability of Dysphonia Measurements for Telemonitoring of Parkinson’s Disease. *IEEE T Bio-Med Eng*, 2009, ročník 56, č. 4, s. 1015–1022.
- [19] LITTLE, M., McSHARRY, P., MOROZ, I., ROBERTS, S. Nonlinear, Biophysically-Informed Speech Pathology Detection. In *2006 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2006. ICASSP 2006 Proceedings. Toulouse, France. s. II-1080- II-1083.
- [20] LITTLE, M., MCSHARRY, P. E., ROBERTS, S. J., COSTELLO, D. A. E., MOROZ, I. M. Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. *Biomed Eng Online*, ročník 6, 2007, str. 23, ISSN 1475-925X.

- [21] MEKYSKA, J., SMÉKAL, Z., KOŠTÁLOVÁ, M., MRAČKOVÁ, M., SKUTILOVÁ, S., REKTOROVÁ, I. Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení, *Cesk Slov Neurol N*, 2011, ročník 74/107(6).
- [22] METTER, J., HANSON, W. Clinical and acoustical variability in hypokinetic dysarthria. *Journal of Communication Disorders*, ročník 1986, s. 347–366.
- [23] PENG, H., LONG, F., DING, C. Feature Selection Based on Mutual Information: Criteria of Max-dependency, Max-relevance, and Min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, ročník 27, č. 8, s. 1226–1238, Aug. 2005.
- [24] PRATAMA, S.F., MUDA, A.K., CHOO, Y., MUDA, N.A. Computationally Inexpensive Sequential Forward Floating Selection for Acquiring Significant Features for Authorship Invarianceness in Writer Identification. *International Journal on New Computer Architectures and Their Applications*, s. 581–598, 2011, ISSN: 2220-9085.
- [25] PSUTKA, J., MÜLLER, L., MATOUŠEK, J., RADOVÁ, V. *Mluvíme s počítačem česky*. 1. vyd. Praha: Academia, 2006. ISBN 80-200-1309-1.
- [26] PUDIL, P., NOVOVIČOVÁ, J., KITTLER, J. Floating search methods in feature selection. *Pattern Recogn. Lett.* 15, 11 (November 1994), 1119–1125.
- [27] REKTOROVÁ, I. Současné možnosti diagnostiky a terapie Parkinsonovy nemoci. *Neurol. praxi*, 2009, ročník 10 (Supl 2), s. 5–36.
- [28] RUSZ, J., CMEJLA, R., TYKALOVA, T. RUZICKOVÁ, H., KLEMPÍR, J. aj. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *J Acoust Soc Am*, ročník 134, č. 3, 2013, s. 2171–2181.
- [29] SANGKIL L., JIEUN, K., INSUNG, L. Speech/Audio Signal Classification Using Spectral Flux Pattern Recognition. *Signal Processing Systems (SiPS), 2012 IEEE Workshop on*, s. 232–236, 17–19 Oct. 2012.
- [30] SAPIR, S., RAMIG, L., FOX, C. Speech and Swallowing Disorders in Parkinson Disease. *Curr Opin Otolaryngol Head Neck Surg*, 2008, ročník 16, č. 3, s. 205–210.
- [31] SAPIR, S., RAMIG, L. O., SPIELMAN, J. L., FOX, C. Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech. *J Speech Lang Hear Res*, ročník 53, č. 1, 2010, str. 1–20.

- [32] SKODDA, S., VISSER, W., SCHLEGEL, U. Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *J Neural Transm*, ročník 117, 2010, s. 197–205.
- [33] SMÉKAL, Z. *Číslíkové zpracování řeči*. Skriptum VUT FEKT v Brně, 2013. 171 s.
- [34] SMÉKAL, Z., MEKYSKA, J., GALAZ, Z., MZOUREK, Z., REKTOROVÁ, I., FAUNDEZ-ZANUY, M. Analysis of Phonation in Patients with Parkinson's Disease using Empirical Mode Decomposition.
- [35] SONG, Y., WANG, W. H., GUO, F. J. Feature extraction and classification for audio information in news video. In *Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009. International Conference on*, 2009, s. 43–46.
- [36] TSANAS, A., LITTLE, M., McSHARRY, P., RAMIG, L. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the Royal Society Interface*, ročník 8, č. 59, 2010, s. 842–855.
- [37] TSANAS, A., LITTLE, M., McSHARRY, P., SPIELMAN, J., RAMIG, L. Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease. *IEEE Trans Biomed Eng*, ročník 59, č. 5, 2012, s. 1264–1271.
- [38] VAZIRI, G., ALMASGANJ, F., BEHROOZMAND, R. Pathological assessment of patients' speech signals using nonlinear dynamical analysis. *Comput Biol Med*, ročník 40, č. 1, 2010, s. 54–63, ISSN 0010-4825.
- [39] ZAMIŠKOVÁ, G., RESSNER, P., DLOUHÁ, J., ŠIGUTOVÁ, D. Poruchy řeči u Parkinsonovy nemoci, *Neurol. praxi*, 2010, č. 11(2), s. 112–116.

# SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

$\alpha$	Hladina významnosti
$\rho_{\text{Sp}}$	Spearmanův koeficient pořadové korelace
$\Delta A$	Rozdíl maximální a minimální hodnoty $A$
$A$	Amplituda hlasivkových pulsů
ACC	Accuracy – přesnost klasifikace
AE	Approximate Entropy – Aproximační entropie
AE GK	Approximate Entropy with Gaussian kernel – aproximační entropie s Gaussovským jádrem
ANOVA	Analysis of variance
APQ3	three-point Amplitude Perturbation Quotient
APQ5	five-point Amplitude Perturbation Quotient
APQ11	11-point Amplitude Perturbation Quotient
CD	Correlation Dimension – korelační dimenze
CHNR	Cepstral Harmonic-to-Noise Ratio
CPP	Cepstral Peak Prominence
DDK	Diadochokinetic Tasks – Diadochokinetické úlohy
DFA	Detrended Fluctuation Analysis
$\Delta E$	Rozdíl maximální a minimální hodnoty $E$
$E$	Energie signálu
EMD-ER	Empirical Mode Decomposition Excitation Ratio
$F_0$	Kmitočet základního tónu
$F_n$	Formantový kmitočet
$FB_n$	Šířka pásma $n$ -tého formantu

FCR	Formant Centralization Ratio – centralizační poměr formantů
FD	Fractal Dimension – fraktální dimenze
FFT	Fast Fourier Transform – rychlá Fourierova transformace
FLUF	Fraction of Locally Unvoiced Frames – podíl neznělých segmentů
FMMI	First Minimum of Mutual Information function – minimum funkce vzájemné informace
FN	False Negative
FP	False Positive
FR	Formant Ratio – poměr formantů $F_{2i}/F_{2u}$
$F_0$ VR	$F_0$ Variation range
GNE	Glottal to Noise Excitation
GQ	Glottal Quotients – hlasivkové kvocienty
$GQ_{\text{closed}}$	Směrodatná odchylka doby, kdy jsou hlasivky uzavřené
$GQ_{\text{open}}$	Směrodatná odchylka doby, kdy jsou hlasivky oteřené
$H_0$	Nulová hypotéza
$H_1$	Alternativní hypotéza
$H(X)$	Entropie náhodné veličiny $X$
$H(X, Y)$	Sdružená entropie $X$ a $Y$
HD	Hypokinetická dysartrie
HE	Hurst Exponent – Hurstův exponent
HNR	Harmonic-to-Noise Ratio
HZ-CRR	High Zero-Crossing Rate Ratio
ir	Interpercentile range – mezipercentilový rozsah
KNN	$k$ -Nearest Neighbour – $k$ -nejbližších sousedů
LASSO	Least Absolute Shrinkage and Selection Operator



LLBFS	Local Learning-Based Feature Selection
LLE	Largest Lyapunov Exponent – největší Lyapunovův exponent
lnVSA	logarithmic Vowel Space Area – logaritmická oblast hlasového traktu
LPC	Linear Prediction Coefficient – lineární predikční koeficient
MFCC	Mel Frequency Cepstral Coefficients – Melovské keprální koeficienty
MI	Mutual Information – vzájemná informace
MMSE	Mini-Mental State Examination – test kognitivních funkcí
mRMR	Minimum Redundancy Maximum Relevance
NHR	Noise-to-Harmonic Ratio
NNE	Normalized Noise Energy
No.	Počet vybraných příznaků
$p_{MW}$	Hladina významnosti (Mann-Whitneyův U test)
$p_{Sp}$	Hladina významnosti (Spearmanův korelační koeficient)
PARCZ	Czech Parkinsonian Speech Database
PCA	Principal Components Analysis
PECM	Pitch Energy Cepstral Measure
PN	Parkinsonova nemoc
PPE	Pitch Period Entropy
PPQ5	five-point Pitch Perturbation Quotient
$Q_i$	Pořadí náhodné veličiny
$R_i$	Pořadí náhodné veličiny
$r_S$	Odhad Spearmanova koeficientu pořadové korelace
RAP	Relative Average Perturbation
RE	Rényi Entropy – Rényiho entropie
rel $F_0$ SD	Relativní směrodatná odchylka $F_0$

rel $F_0$ VR	Relativní kolísání $F_0$
RF	Random Forests – náhodné stromy
RPDE	Recurrence Probability Density Entropy
RR	Rahmonic Ratio
SA	Sample Entropy – vzorková entropie
SAGK	Sample Entropy with Gaussian kernel – vzorková entropie s Gaussovským jádrem
SDBM	Spectral Distance Based on Module
SDBP	Spectral Distance Based on Phase
SE	Shannon Entropy – Shannonova entropie
SEN	Sensitivity – sensitivita
SF	Spectral Flux – spektrální tok
SFFS	Sequential Forward Floating Search – sekvenční dopředná selekce příznaků
SPE	Specificity – specifická
SPI	Soft Phonation Index
STD	Standard Deviation – směrodatná odchylka
SVM	Support Vector Machines
TKEO	Teager-Kaiser Energy Operator – Teagerův-Kaiserův energetický operátor
TN	True Negative
TP	True Positive
TSS	Trade-off between sensitivity and specificity – kompromis mezi sensitivitou a specifickou
UPDRS	Unified Parkinson’s Disease Rating Scale – unifikovaná škála pro hodnocení Parkinsonovy nemoci
VAI	Vowel Articulation Index – artikulační index samohlásek

VFER	Vocal Fold Excitation Ratio
VSA	Vowel Space Area – oblast hlasového traktu
VTI	Voice Turbulence Index
ZCR	Zero-Crossing Rate – počet průchodů signálu nulovou úrovní
ZL	Ziv-Lempel

# SEZNAM PŘÍLOH

A	Obsah CD	61
B	Obsah adresáře feature_toolbox	62

## A OBSAH CD

Na přiloženém CD se nachází:

- diplomová práce ve formátu pdf,
- složka se zdrojovými soubory `feature_toolbox`.

Databáze PARCZ není součástí CD z důvodu autorských práv Fakultní nemocnice u sv. Anny v Brně. Kód programu byl vytvořen a testován v programu Matlab ve verzi R2011b.

## B OBSAH ADRESÁŘE FEATURE\_TOOLBOX

Adresář `feature_toolbox` obsahuje všechny součásti systému vytvořené a použité v rámci diplomové práce. Je rozčleněn do následujících složek:

- `cpp` – obsahuje spustitelný soubor `cpps.exe` nutný pro výpočet řečového příznaku CPP [12];
- `DFA` – soubory pro výpočet DFA<sup>1</sup> [19];
- `feature_func` – funkce `.m` pro výpočty jednotlivých řečových příznaků (některé z toolboxu Voice Analysis Toolbox<sup>2</sup> [36]);
- `feature_matrix_func` – funkce `.m` pro výpočet a práci s maticí příznaků, algoritmus SFFS<sup>3</sup>;
- `mRMR` – soubory algoritmu mRMR<sup>4</sup> [23];
- `praatcon` – konzolová verze programu Praat s obslužnými skripty [4];
- `RPDE` – soubory pro výpočet RPDE<sup>1</sup> [20];
- `voicebox` – funkce `.m` z toolboxu Voicebox [5].

Program je spustitelný dvojicí skriptů:

- `individual_analysis.m` – provede výpočet matice příznaků a následnou analýzu popsanou v kapitole 5.1;
- `feature_selection.m` – z matice příznaků dle zvoleného scénáře provede výběr příznaků algoritmem SFFS dle kapitoly 5.2.

Zdrojové soubory jsou opatřeny komentáři, které podrobně popisují činnost jednotlivých skriptů a funkcí včetně jejich parametrů.

---

<sup>1</sup>Dostupné z <http://www.maxlittle.net/software>

<sup>2</sup>Dostupné z <https://people.maths.ox.ac.uk/tsanas/software.html>

<sup>3</sup>Dostupné z <http://users.spa.aalto.fi/jpohjala/featureselection>

<sup>4</sup>Dostupné z <http://penglab.janelia.org/proj/mRMR/#matlab>